运筹与管理科学丛书 4

组合预测方法有效性理论及其应用

陈华友 著



科学出版社
www.sciencep.com

(O-2947.0101)



ISBN 978-7-03-020213-0

销售分类建议：高等数学，管理科学

定　价：45.00元

运筹与管理科学丛书 4

组合预测方法有效性理论及其应用

陈华友 著

科学出版社

北 京

内 容 简 介

本书研究预测方法的有效性理论及其应用，建立了基于不同准则的组合预测模型. 在模型的构造方面，建立了基于预测有效度准则的最优组合预测模型、基于多种诱导有序加权平均算子的最优化组合预测模型、基于相关性指标的最优组合预测模型、基于非线性加权平均的最优组合预测模型等；在模型的有效性理论的探讨方面，针对多种准则下最优组合预测提出了优性组合预测、预测方法优超和冗余度等概念，给出了非劣性组合预测、优性组合预测、冗余预测方法的存在性以及冗余预测信息的判定；在模型的应用方面，探讨了组合预测技术在证券组合投资、剩余劳动力优化配置、组合赋权决策等领域的应用.

本书可作为高等院校应用数学、运筹学、统计学、管理科学和系统工程专业的高年级本科生和研究生教材，也可作为工程技术人员、管理干部和相关学者的参考书.

图书在版编目（CIP）数据

组合预测方法有效性理论及其应用 / 陈华友著. —北京：科学出版社，2008
（运筹与管理科学丛书；4）
ISBN 978-7-03-020213-0

Ⅰ. 组… Ⅱ. 陈… Ⅲ. 组合数学-数学预测 Ⅳ. O22

中国版本图书馆 CIP 数据核字（2007）第 196409 号

责任编辑：赵彦超 / 责任校对：陈玉凤
责任印制：赵德静 / 封面设计：王 浩

科 学 出 版 社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com
新 蕾 印 刷 厂 印刷
科学出版社发行 各地新华书店经销
*
2008 年 2 月第 一 版 开本：B5(720×1000)
2008 年 2 月第一次印刷 印张：17
印数：1—3 000 字数：315 000

定价：45.00 元
（如有印装质量问题，我社负责调换〈环伟〉）

《运筹与管理科学丛书》序

运筹学是运用数学方法来刻画、分析以及求解决策问题的科学. 运筹学的例子在我国古已有之，春秋战国时期著名军事家孙膑为田忌赛马所设计的排序就是一个很好的代表. 运筹的重要性同样在很早就被人们所认识，汉高祖刘邦在称赞张良时就说道："运筹帷幄之中，决胜千里之外."

运筹学作为一门学科兴起于第二次世界大战期间，源于对军事行动的研究. 运筹学的英文名字 Operational Research，诞生于 1937 年. 运筹学发展迅速，目前已有众多的分支，如线性规划、非线性规划、整数规划、网络规划、图论、组合优化、非光滑优化、锥优化、多目标规划、动态规划、随机规划、决策分析、排队论、对策论、物流、风险管理等.

我国的运筹学研究始于 20 世纪 50 年代，经过半个世纪的发展，运筹学队伍已具相当大的规模. 运筹学的理论和方法在国防、经济、金融、工程、管理等许多重要领域有着广泛应用，运筹学成果的应用也常常能带来巨大的经济和社会效益. 由于在我国经济快速增长的过程中涌现出了大量迫切需要解决的运筹学问题，因而进一步提高我国运筹学的研究水平、促进运筹学成果的应用和转化、加快运筹学领域优秀青年人才的培养是当今我们面临的十分重要、光荣、同时也是十分艰巨的任务. 我相信，《运筹与管理科学丛书》能在这些方面有所作为.

《运筹与管理科学丛书》可作为运筹学、管理科学、应用数学、系统科学、计算机科学等有关专业的高校师生、科研人员、工程技术人员的参考书，同时也可作为相关专业的高年级本科生和研究生的教材或教学参考书. 希望该丛书能越办越好，为我国运筹学和管理科学的发展做出贡献.

袁亚湘

2007 年 9 月

序 言

复杂的社会经济系统面临一个变化迅速的外界环境．由于预测的不确定性因素显著增加，利用传统的单个预测模型进行预测的缺陷表现为信息源的不广泛性以及对模型设定形式的敏感性等．因此，采用单个预测模型进行预测往往难以达到较为理想的预测效果．为了分散预测的风险，组合预测方法应运而生．

组合预测方法有效性理论的研究已经引起学术界的关注．本书建立了基于不同准则的组合预测模型．在模型的构造方面，建立了基于预测有效度准则的最优组合预测模型、基于多种诱导有序加权平均算子的最优化组合预测模型、基于相关性指标的最优组合预测模型、基于非线性加权平均的最优组合预测模型等；在模型的有效性理论探讨方面，针对多种准则下最优组合预测提出了优性组合预测、预测方法优超和冗余度等概念，给出了非劣性组合预测、优性组合预测、冗余预测方法的存在性以及冗余预测信息的判定；在模型的应用方面，探讨了组合预测技术在证券组合投资、剩余劳动力优化配置、组合赋权决策等领域的应用．作者在写作过程中注意到以下两个方面：一是力求理论上的系统性和新颖性；二是考虑到组合预测方法具有良好的应用背景，增加了一些实证分析的内容，尽量做到理论和应用的统一．

自 1998 年以来，作者一直从事组合预测方法的理论和应用研究．2005 年，作者主持了国家自然科学基金项目“诱导有序加权平均组合预测模型的构建及其有效性理论和应用研究” (70571001)；2007 年，作者主持了安徽省自然科学基金项目“基于多种不确定型偏好信息及其集结算子的群决策模型方法”．本书是这两个项目研究的一个阶段性成果．同时，本书也是作者多年来在组合预测和决策分析方面的理论研究和应用研究成果的一个系统总结．作者感谢国家自然科学基金委员会和安徽省自然科学基金委员会对本书的资助．作者是安徽大学数学创新团队的重要成员，也要感谢安徽大学创新团队对本书的资助．

本书可作为高等学校应用数学、运筹学、统计学、管理科学和系统工程专业的高年级本科生和研究生教材，也可作为工程技术人员、管理干部和相关学者的参考书．

在本书的编辑出版过程中，得到了科学出版社编辑的大力支持和帮助，作者在此表示衷心的感谢．

最后，向多年来关心支持我的工作的领导、老师、同事和家庭表示衷心的感谢．

由于本人的学识水平有限，书中难免存在缺点和错误，欢迎读者不吝赐教．

陈华友
2007 年 10 月于合肥

目　录

第1章 绪　论

1.1 预测的基本概念及其遵循的基本原则

预测是为适应社会经济的发展与管理的需要而产生和发展起来的. 预测作为一种社会实践活动, 已有几千年的历史. 可以说, 自人类有文明历史以来, 就存在预测活动. 古代的预测与占卜术密切相关, 是在先兆、经验的前提下推测未来. 由于当时人们对客观物质世界的认识水平较低, 古代的预测多半具有唯心主义和封建迷信的色彩.

随着人类社会的发展, 社会生产力得到了较大提高. 特别是作为生产力第一要素的科学技术水平的显著提高, 科学预测逐步取代了迷信预测和经验预测而发展成为一门学科. 预测学就是总结预测活动的经验所形成的理论概括, 它产生的历史并不长. 预测真正成为一门自成体系的独立的学科仅仅是近几十年的事情. 特别是二次世界大战以后, 由于科学技术和世界经济取得了前所未有的快速发展, 社会经济现象的不确定因素显著增加, 诸如政治危机、经济危机、能源危机、恐怖活动等. 所有这些不确定因素增加了人们从心理上了解和掌握未来的必要性和迫切性. 人们日益意识到科学预测的重要性, 这也就成为预测学进一步发展的推动力.

对于“预测”一词, 可以从不同的角度来理解. 它有三个含义, 即预测工作、预测结果、预测学.

首先, 从预测工作来看, 它是指一种实践活动. 预测是根据不确定事件或未知事件的过去和现状的信息来推知、估计未来, 探索事件发展变化的规律. 即根据已知推断未知的过程.

其次, 从预测结果来看, 它是预测工作的成果和“产品”. 具体表现为预测工作过程所获得的预测值. 这些预测值反映社会经济现象的数量特征及其规律性.

再次, 从预测学来看, 它是阐述预测方法的一门学科和理论. 科学预测方法是采用科学的判断和计量方法, 对未来事件的可能变化情况作出事先推测的一种技术. 预测学是一门应用方法论的学科. 科学预测方法要求根据社会经济现象的历史和现实, 综合多方面的信息, 运用定性和定量相结合的分析方法, 揭示客观事物的发展变化规律, 并指出事物之间的联系、未来发展的途径和结果等.

上述三个含义既有区别也有联系. 预测结果是预测工作的成果. 预测学是预测工作的理论概括和总结, 因此, 预测学阐述的预测方法对预测工作起着指导作用. 预测工作一方面接受预测方法对其的指导作用, 另一方面可以用来检验预测理论方法

的正确性, 从而促进预测理论方法的发展. 预测学与预测工作、预测结果之间的关系表明理论来源于实践, 又反过来服务于实践, 体现着理论与实践的辩证关系.

预测学突破了自然科学和社会科学的界限, 已发展成为一门综合性的学科, 目前预测学应用研究有很大的拓展, 广泛应用于人口、环境、资源、教育、金融、交通运输、城市规划、医药卫生、材料科学、科技管理等领域. 可见, 预测方法与各个学科、各个部门均有密切联系. 同时, 预测学理论研究有了新的进展, 但是我们还不能说预测学已经发展得很成熟. 它在以较快的速度继续向前发展. 在发展过程中不断地吸收其他学科的营养, 进一步丰富和完善自己.

一般而言, 预测遵循以下基本原则 [1]:

(1) 连贯性原则. 预测对象具有的规律性不仅在过去和现在起作用, 而且在未来的一段时间内继续发挥作用, 这种连贯性包括时间的连贯性和预测系统结构的连贯性.

(2) 相关类推原则. 预测对象的发展变化与某些因素密切. 有的呈正相关关系, 有的呈负相关关系. 因此, 类推原则要求在建立适当的预测模型后, 根据相关因素发展变化来类推预测对象的规律.

(3) 概率性原则. 预测对象的发展既受到偶然因素的影响, 又受到必然因素的影响. 概率性原则要求利用统计方法可以获得预测对象发展的必然规律.

1.2　对传统预测方法的评价

预测根据其目标和特点不同, 可以分成不同的类别. 传统的预测方法按属性不同, 可以分为定性预测方法和定量预测方法.

定性预测方法就是以人的经验、事理等主观判断为主的预测方法, 对事物未来的性质作出描述. 一般地说, 定性预测方法适用于缺少历史统计资料, 而更多地需要依赖专家的经验的情况下使用. 定性预测方法通常有德尔菲法、主观概率法、市场调查法、领先指标法、模拟推理法和相关因素分析法等. 定性预测法的特点可归纳为：

(1) 强调对事物发展的性质进行描述性的预测. 这主要通过专家的经验以及分析判断能力, 尤其是在对预测对象所掌握的历史数据不多或影响预测对象因素众多、复杂, 难以作出定量分析的情况下, 定性预测方法是较好的可行方法.

(2) 强调对事物发展的趋势、方向和重大转折点进行预测. 例如, 某商品在市场上所处的阶段、市场总体形势的变化、国家产业政策的变化、新产品的开发、企业经营环境分析等.

从上面定性预测方法的特点可知, 定性预测法的优点在于预测事物未来发展性质方面, 且定性预测法的灵活性较强, 能充分发挥人们的主观能动性. 同时, 定性预

测法预测简单迅速, 可节省一定的人力、物力和财力. 当然, 定性预测方法也存在缺点. 表现为它受人们的主观因素的影响较大. 这是因为定性预测方法主要依赖于人们的知识、经验和能力的大小等, 因此缺乏成套的数学模型, 难以对事物发展作出数量上的精确度量.

定量预测方法就是利用预测对象的历史和现状的数据, 按变量之间的函数关系建立数学模型, 从而计算出预测对象的预测值. 显然, 定量预测方法适用于历史统计资料较为丰富的情况. 定量预测方法通常有移动平均法、指数平滑法、线性回归法、非线性回归法、马尔可夫预测法、投入产出预测法、灰色预测法、Box-Jenkins 模型法、经济计量模型法、干预分析模型法等. 定量预测方法的特点可归纳如下:

(1) 强调对事物发展的数量方面进行较为精确的预测. 这主要通过历史统计数据建立相应的数学模型, 对事物发展作出数量上的预测.

(2) 强调对事物发展的历史统计资料利用的重要性. 目前, 国民经济核算体系及其他统计数据正好为定量预测法提供了信息来源.

(3) 强调建立数学模型的重要性, 且要利用电子计算机来解决定量预测法中复杂的数学模型的参数计算问题. 目前, 电子计算机的迅速发展和普及, 为定量预测法提供了良好的技术条件.

从上面定量预测方法的特点可知, 定量预测法的优点偏重于预测事物未来发展数量方面的准确描述. 它较少依赖于人的知识、经验等主观因素, 而是更多地依赖于预测对象客观的历史统计资料, 利用电子计算机对数学模型进行大量的计算而获得预测结果. 其缺点是: 对预测者的素质要求较高, 预测者必须掌握数学方法、计算机技术及相应的专门理论; 另外, 定量预测法的精确度较多地依赖于统计资料的质量和数量. 同时, 若预测对象的系统结构发生质的变化, 相应的统计数据发生较大的波动, 此时定量预测法难以获得满意的预测结果.

1.3　组合预测方法研究的现状

实际的预测对象可能是较为复杂的社会经济系统, 有多种错综复杂的因素对其产生影响. 有些是基本因素, 有些是偶然因素. 预测者常常在不同的假设条件下, 采用不同的单项预测方法对同一预测问题建立多种预测模型, 然后按照统计假设检验, 从众多的预测方法中选择结果最好的一个, 而排除了其他单项预测方法. 这是不是提高预测精度的最佳的办法呢? 回答是否定的.

显然, 不同的定性预测模型方法和定量预测模型方法各有其优点和缺点, 它们之间并不是相互排斥, 而是相互联系、相互补充. 由于每种预测方法利用的数据不尽相同, 不同的数据从不同的角度提供各方面有用的信息. 在预测的过程中, 如果想当然的认为某个单项预测方法的预测误差较大, 就把该种预测方法弃之不用, 这

可能造成部分有用的信息丢失. 因此, Bates 和 Granger[2] 于 1969 年首次提出一个合理的做法, 即综合考虑各单项预测方法的特点, 将不同的单项预测方法进行组合, 提出组合预测方法的概念. 也就是说, 即使一个预测误差较大的预测方法, 如果它包含系统独立的信息, 当它与一个预测误差较小的预测方法组合后, 完全有可能增加系统的预测性能.

若预测者只用一种预测方法进行预测, 则这种预测方法的选择是否适当就显得很重要. 如果预测者选择预测方法不当, 就可能要冒一定决策失误的风险. 在预测实践中, 若把多种单项预测方法正确地结合起来使用, 则会使得组合预测结果对某单个较差的预测方法不太敏感. 因此, 组合预测一般能提高预测的精确度和可靠度.

所谓组合预测就是设法把不同的预测模型组合起来, 综合利用各种预测方法所提供的信息, 以适当的加权平均形式得出组合预测模型. 组合预测最关心的问题就是如何求出加权平均系数, 使得组合预测模型更加有效地提高预测精度. 组合预测在国外称为 combination forecasting 或 combined forecasting, 在国内也称为综合预测等.

1969 年, Bates 和 Granger[2] 首次对组合预测方法进行系统研究. 其研究成果引起预测学者的重视. 20 世纪 70 年代以后, 组合预测方法的研究进一步得到了重视. 1989 年, 国际预测领域的权威学术刊物 *Journal of Forecasting* 还出版了组合预测方法专辑 [3~13]. 这充分说明了组合预测方法在预测学中的重要地位 [14].

最近十几年, 国内预测学界也非常重视组合预测方法的研究, 取得一系列的研究成果. 唐小我教授的研究成果尤其突出 [14~22,26]. 组合预测方法的研究成果主要发表在《管理科学学报》[21]、《系统工程理论与实践》[24,25]、《电子科技大学学报》[18]、《预测》[16]、《系统工程学报》[37]、《系统工程理论方法应用》[27]、《控制与决策》[23] 等学术刊物上.

以上各种不同的组合预测方法中, 实际应用和理论研究最多的是最小方差方法, 且大体上以绝对误差作为准则来计算出组合预测方法的权系数向量. 但是, 评价预测精度的指标除了绝对误差以外, 还可以用相对误差指标或预测精度指标来反映. 因此, 目前组合预测方法的研究并不完善, 需要进一步加强研究.

对国内外文献的系统分析表明, 相关研究呈现以下特点:

(1) 提出多种准则下的组合预测模型, 对组合预测模型的求解和有效性的实证研究较为深入, 但缺乏多种准则框架下的组合预测模型有效性的理论研究成果. 目前, 国内外学者主要提出以下一些组合预测方法: 最小方差方法 [2,28~31,36]、无约束最小二乘方法 [32]、约束最小二乘方法 [33,34]、Bayes 方法 [35,38]、基于不同准则和范数的组合预测方法 [38,39] 、递归组合预测方法 [40,41] 等.

(2) 对组合预测方法的有效性的理论研究已经引起学术界的关注. 文献 [40] 针对无非负约束的以误差平方和达到最小的组合预测模型提出了优性组合预测的概

念, 并利用组合预测绝对误差信息矩阵的性质判断简单平均方法是优性组合预测的条件; 文献 [42] 研究了该模型组合预测误差的界; 文献 [72] 提出了基于预测有效度的组合预测模型, 并给出组合预测权系数的线性规划的求解方法; 文献 [73] 针对此模型探讨其有效性; 文献 [74] 研究了考虑预测精度标准差的预测有效度的组合预测模型的性质. 但是, 总体来说, 针对其他准则[43] 下的组合预测模型及加权调和平均组合预测模型[44], 加权几何平均组合预测[45] 的有效性的理论研究成果尚需进一步研究.

(3) 现有的组合预测方法根据单项预测方法的种类不同而赋予不同的加权平均系数, 其特点是同一个单项预测方法在各个时点的加权平均系数是相同的. 然而, 实际的情况是同一个单项预测方法在不同时刻的表现并不相同, 即在某个时点上预测精度较高, 而在另一时点上预测精度较低, 有必要讨论有序加权几何平均的新的组合预测模型, 新模型赋权的基本思想是依据每个单项预测方法在各个时点的拟合精度的高低进行有序赋权.

1.4 本书的主要内容

本书的主要内容包括 10 章, 具体安排如下:

第 1 章是绪论. 首先叙述了预测的基本概念, 然后分析了定性预测方法和定量预测方法的特点, 指出了两种预测方法的信息互补性, 说明了组合预测方法产生的背景, 最后综述了组合预测方法研究的国内外现状, 指出了进一步加强研究的必要性.

第 2 章是常用的单项预测模型. 主要介绍常用的单项预测模型的预测方法, 包括时间序列预测模型、回归分析预测模型、随机时间序列预测模型、灰色预测 GM(1, 1) 模型、季节变动时间序列的预测模型.

第 3 章是非最优组合预测模型研究. 由于在求解最优组合预测模型时, 最优组合预测的权系数有可能出现负的情况, 而负的组合预测的权系数没有实际的意义. 所以通常有两种解决问题的办法: 一是在最优组合预测模型中增加非负约束条件; 二是研究满足非负约束的非最优组合预测模型. 因此, 本章进一步探讨确定组合预测的权系数的非最优组合预测模型, 以满足决策者的需求.

第 4 章是基于预测误差指标的最优组合预测模型. 目前, 国内外学者提出各种不同的组合预测方法主要是以绝对误差作为准则来计算出组合预测方法的权系数向量. 本章介绍了几种主要的组合预测模型、包括以预测误差平方和达到最小的线性组合预测模型、以离差绝对值和达到最小的线性组合预测模型、以最大离差绝对值达到最小的线性组合预测模型、以组合预测误差全距作为目标函数的组合预测模型、非负可变加权系数的组合预测模型, 最后给出基于预测误差指标的最优组合

预测模型的实例分析.

第 5 章是基于预测有效度的最优组合预测的有效性理论. 目前组合预测模型大多研究的是关于最小方差方法的模型解的求法、组合结构特征以及预测误差平方和的取值范围 [56,59] 等问题. 因此, 本章重点提出了基于预测有效度组合预测模型, 针对基于一阶或二阶预测有效度的组合预测模型, 提出了若干新概念, 并对其几个基本问题进行研究. 这些基本问题包括非负权重最优组合预测方法的预测有效度是否一定大于各个单项预测方法预测有效度中的最大者; 当参加组合的预测方法增多时, 非负权重最优组合预测方法的预测有效度是否一定增大; 最优组合预测方法是否存在冗余预测方法以及冗余信息的判定等. 最后对回归型组合预测模型进行权系数估计及其显著性检验.

第 6 章是非线性加权平均的最优组合预测模型的有效性理论. 本章提出了基于 L_2 和 L_1 范数的加权几何平均组合预测方法, 给出模型的求解方法, 并进行了实例分析. 同时, 针对基于 L_1 范数的加权几何平均组合预测方法, 调和平均的组合预测方法的性质进行了研究, 并对广义加权算术平均组合预测法的最优化理论基础及性质进行了探讨.

第 7 章是基于相关性指标的最优组合预测模型的有效性理论. 基于相关性指标的最优组合预测模型是研究组合预测方法的一个新途径. 本章针对四种基于相关性指标的最优组合预测模型, 即基于关联度最大化组合预测模型、基于相关系数最大化组合预测模型、基于夹角余弦最大化组合预测模型、基于 Theil 不等系数最小化组合预测模型, 研究了它们的基本结构特征. 在提出新的优性组合预测、预测方法优超、冗余度等概念的基础上, 探讨了非劣性组合预测、优性组合预测以及冗余预测方法的存在性, 并给出冗余信息的判定定理. 实例分析表明, 这些方法有较大的实际应用价值.

第 8 章是基于诱导有序信息集结算子的最优组合预测模型及其有效性理论. 组合预测方法之所以能提高预测精度, 是因为它能有效地提取各单个的预测模型提供的有用信息, 因此, 组合预测方法的实质也是一种信息的有效集成方法. 有序加权平均算子和诱导有序加权平均算子均是介于最大算子与最小算子之间的一种信息集成方法, 常规的加权算术平均算子是其特例. 本章探讨了有序加权平均算子和诱导有序加权平均算子的概念和性质. 在此基础上提出了基于 IOWA 算子的组合预测方法、基于 IOWGA 算子的组合预测方法和基于 IOWHA 算子的组合预测方法, 给出了相应的权系数的确定的数学规划方法. 实例分析结果表明, 模型能有效提高组合预测精度. 本章最后探讨了一类 OWA 算子的组合预测模型的有效性理论.

第 9 章是组合预测技术的应用研究. 从信息的利用角度来看, 组合预测综合利用各种单项预测方法所提供的信息, 以适当的加权平均形式得出组合预测模型. 因此, 组合预测方法可以降低信息的不完备性, 从而使得组合预测模型更加有效地提

高预测精度. 本章把组合预测技术的基本思想移植到其他领域, 扩大其应用范围. 本章探讨了组合预测技术在多属性决策 (或多指标评价)、证券组合投资、剩余劳动力配置等方面的应用.

第 10 章是组合判断矩阵及相关决策问题研究. 由于决策问题的复杂性, 决策者很难直接提供方案的排序结果, 而是给出方案两两比较的判断矩阵的偏好信息. 目前从判断矩阵中元素表现的方法来划分, 可分为互反判断矩阵和模糊 (或互补) 判断矩阵两大类. 有关互反判断矩阵的研究已趋于完善, 而模糊判断矩阵尚在研究之中. 本章研究组合判断矩阵的相容性与一致性关系, 根据模糊判断矩阵互补性的特点, 提出一个相容性的新指标, 并研究模糊判断矩阵相容性和一致性的关系, 在相容性概念的基础上, 研究了加权几何平均组合判断矩阵的相容性以及相容性和一致性的关系. 在相对较弱的条件下, 获得了组合判断矩阵不仅与其自身的特征矩阵具有满意的相容性, 而且它还与加权几何平均组合排序向量所构成的特征矩阵具有满意的相容性. 这为在群组决策中使用加权几何平均向量排序法提供理论依据. 另外还探讨了模糊判断矩阵的相容性研究、模糊判断矩阵排序的最小偏差法的性质等.

第 2 章　常用的单项预测模型

组合预测方法是以单项预测方法作为基础的, 因此, 本章介绍几种常用的单项预测方法的基本原理[46~48], 包括移动平均法、指数平滑法、线性回归法、Box-Jenkins 模型法、灰色预测模型法等. 这些单项预测方法利用预测对象的历史和现状的数据, 建立相应的数学模型, 从而计算出预测对象的预测值.

2.1　时间序列预测模型

时间序列是指社会经济现象或自然现象的某种指标按时间先后顺序所排列的数字序列, 它包含两个要素: 一个是现象所属的时间, 另一个是其对应的指标值. 一般的时间序列可表示为$\{x_t, t=1,2,\cdots,n\}$.

2.1.1　具有局部水平趋势的平滑预测模型

设有时间序列$\{x_t, t=1,2,\cdots,n\}$, 若它在某一局部时间范围内的指标值围绕某一水平线上下波动, 而在较长时间范围内的指标值是缓慢移动的, 则称该时间序列具有局部水平趋势. 现介绍三种预测模型.

1. 简单移动平均预测法

所谓简单移动平均法就是从时间序列的第一项数值开始, 按一定的项数求时间序列的序时平均数, 然后逐项移动, 边移动边平均. 这样就可以得到一个由移动平均数构成的新的时间序列. 它是对原有的时间序列的不规则变动进行了修匀. 简单移动平均预测法的计算公式如下

$$\hat{x}_{t+1} = \frac{x_t + x_{t-1} + \cdots + x_{t-N+1}}{N}, \tag{2.1.1}$$

其中 $\hat{x}_{t+1}$ 表示 $t+1$ 时刻时间序列指标的预测值, N 称为移动平均项数.

为简化计算, 由式 (2.1.1) 可得如下递推公式

$$\begin{aligned}\hat{x}_{t+2} &= \frac{x_{t+1} + x_t + \cdots + x_{t-N+2}}{N} \\ &= \frac{(x_{t+1} - x_{t-N+1}) + (x_t + \cdots + x_{t-N+2} + x_{t-N+1})}{N} \\ &= \hat{x}_{t+1} + \frac{x_{t+1} - x_{t-N+1}}{N}.\end{aligned} \tag{2.1.2}$$

在简单移动平均法中, 移动平均项数 N 的选取具有一定的主观性和经验性. 一般可采用取几个不同的 N 值进行试算, 再选取拟合均方误差最小所对应的 N 值.

2. 加权移动平均预测法

在简单移动平均法中, 每个时刻时间序列指标的权系数都是相等的. 然而, 由于近期的统计数据所包含的信息量大, 远期的统计数据所包含的信息量小, 因此, 采用不等权的加权移动平均法更符合实际. 加权移动平均预测法的计算公式如下

$$\hat{x}_{t+1} = \lambda_0 x_t + \lambda_1 x_{t-1} + \cdots + \lambda_{N-1} x_{t-N+1}, \tag{2.1.3}$$

其中 λ_i 为统计数据 x_{t+1-i} 的权系数, 且有 $\sum\limits_{i=0}^{N-1} \lambda_i = 1$, $\lambda_0 \geqslant \lambda_1 \geqslant \cdots \geqslant \lambda_{N-1} \geqslant 0$, $i = 0, 1, \cdots, N-1$.

3. 指数平滑预测法

在加权移动平均法中, 若取移动平均项数 $N = n$, 且权系数 $\lambda_0, \lambda_1, \cdots, \lambda_{n-1}$ 为等比数列, 不妨设 $\lambda_i = kw^i$, $i = 0, 1, \cdots, n-1$, 其中 $w \in (0,\ 1)$, k 为常数. 则称此时的加权移动平均法为指数平滑预测法.

因为权系数满足 $\sum\limits_{i=0}^{n-1} \lambda_i = 1$, 即 $\sum\limits_{i=0}^{n-1} kw^i = 1$, 所以 $k = \dfrac{1-w}{1-w^n}$, 从而指数平滑预测法的计算公式为

$$\hat{x}_{n+1} = \lambda_0 x_n + \lambda_1 x_{n-1} + \cdots + \lambda_{n-1} x_1 = \sum_{i=0}^{n-1} \frac{1-w}{1-w^n} w^i x_{n-i}, \tag{2.1.4}$$

其中 $\hat{x}_{n+1}$ 表示 $n+1$ 时刻时间序列指标的预测值.

因为 $w \in (0,\ 1)$, 所以 $\lim\limits_{n\to+\infty} w^n = 0$, 从而式 (2.1.4) 就转化为

$$\hat{x}_{n+1} = (1-w) \sum_{i=0}^{\infty} w^i x_{n-i}. \tag{2.1.5}$$

若令 $\alpha = 1 - w$, 则 $\alpha \in (0\ ,\ 1)$, $w = 1 - \alpha$, 式 (2.1.5) 就写为

$$\hat{x}_{n+1} = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i x_{n-i}, \tag{2.1.6}$$

其中 α 称为平滑系数. 为简化计算, 由式 (2.1.6) 可得递推公式

$$\begin{aligned}\hat{x}_{n+1} &= \alpha\sum_{i\geqslant 0}(1-\alpha)^i x_{n-i} = \alpha\, x_n + \alpha\sum_{i\geqslant 1}(1-\alpha)^i x_{n-i}\\ &= \alpha\, x_n + (1-\alpha)\left(\alpha\sum_{j\geqslant 0}(1-\alpha)^j x_{n-1-j}\right)\\ &= \alpha\, x_n + (1-\alpha)\hat{x}_n. \end{aligned} \tag{2.1.7}$$

式 (2.1.7) 的直观含义是, 第 $n+1$ 时刻时间序列指标的指数平滑预测值等于第 n 时刻时间序列指标的实际值和第 n 时刻时间序列指标的指数平滑预测值的加权平均. 一般初始的指数平滑预测值可取为初始时刻时间序列的实际值, 即 $\hat{x}_0 = x_1$.

在指数平滑预测法中, 需要讨论平滑系数 α 的选取问题. 常用的选取方法大体上有两种:

(1) 主观法. 根据时间序列指标值的变化情况来确定平滑系数 α. 若时间序列指标值变化缓慢, α 可取小些, 例如 α 在 0.3 左右; 若时间序列指标值变化比较迅速, α 可取大些, 例如 α 在 0.7 至 0.9 左右.

(2) 客观法. 根据指数平滑预测法在拟合区间上误差平方和指标, 使其极小化, 从而求解出最优的 α. 具体来说, 令 $e_t = x_t - \hat{x}_t$, $t = 1, 2, \cdots, n$, 则 e_t 表示第 t 时刻时间序列指标的指数平滑预测误差, 其中 $\hat{x}_t = \alpha\, x_{t-1} + (1-\alpha)\hat{x}_{t-1}$, $\hat{x}_1$ 为给定的第 1 时刻时间序列指标的初始预测值, 显然, e_t 为平滑系数 α 的函数, 设 Q 为指数平滑预测法在拟合区间上误差平方和, 则有

$$Q = \sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n}(x_t - \hat{x}_t)^2.$$

显然, Q 也为 α 的函数, 记为 $Q(\alpha)$, 因此, 最优的 α 选取可表示为如下最优化问题

$$\min_{\alpha\in(0,1)} Q = \sum_{t=1}^{n}(x_t - \hat{x}_t)^2.$$

上述最优化问题的求解可采用一维搜索法.

2.1.2　具有线性趋势的外推预测模型

设时间序列$\{x_t, t=1,2,\cdots, n\}$具有线性趋势, 若它用简单移动平均法进行预测, 则会产生滞后现象. 需要对简单移动平均法和指数平滑预测法进行改进, 为此提出两种预测模型.

1. 二次移动平均预测模型和参数估计

所谓二次移动平均法就是对一次移动平均值序列再进行一次移动平均得到的值, 即

$$\hat{x}_{t+1} = \frac{x_t + x_{t-1} + \cdots + x_{t-N+1}}{N}$$

是一次移动平均值序列, 对其再作一次移动平均, 就得到二次移动平均计算式

$$\hat{\hat{x}}_{t+1} = \frac{\hat{x}_t + \hat{x}_{t-1} + \cdots + \hat{x}_{t-N+1}}{N}, \tag{2.1.8}$$

因为时间序列$\{x_t, t=1,2,\cdots,n\}$具有线性趋势, 设二次移动平均法预测方程为

$$\hat{x}_{t+T} = a_t + b_t T, \tag{2.1.9}$$

其中 a_t, b_t 为待估参数, T 为预测步长. 下面讨论 a_t, b_t 的估计.

实际上, 若时间序列$\{x_t, t=1,2,\cdots,n\}$具有线性趋势方程, 由式 (2.1.9) 得 $t+1$ 时刻的实际值为

$$x_{t+1} = a_t + b_t, \tag{2.1.10}$$

则 $t+1$ 时刻的一次移动平均值为

$$\begin{aligned}\hat{x}_{t+1} &= \frac{x_t + x_{t-1} + \cdots + x_{t-N+1}}{N} = \frac{a_t + (a_t - b_t) + \cdots + [a_t - (N-1)b_t]}{N} \\ &= a_t - \frac{(N-1)b_t}{2}.\end{aligned} \tag{2.1.11}$$

所以由式 (2.1.10) 和式 (2.1.11) 得

$$x_{t+1} - \hat{x}_{t+1} = \frac{N+1}{2} b_t. \tag{2.1.12}$$

同理有

$$\hat{x}_{t+1} - \hat{\hat{x}}_{t+1} = \frac{N+1}{2} b_t. \tag{2.1.13}$$

由式 (2.1.13) 得

$$b_t = \frac{2}{N+1}\left(\hat{x}_{t+1} - \hat{\hat{x}}_{t+1}\right), \tag{2.1.14}$$

式 (2.1.14) 即为 b_t 的估计.

由式 (2.1.12)、(2.1.13) 知 $x_{t+1} = 2\hat{x}_{t+1} - \hat{\hat{x}}_{t+1}$, 由式 (2.1.10)、(2.1.14) 得

$$a_t = x_{t+1} - b_t = 2\hat{x}_{t+1} - \hat{\hat{x}}_{t+1} - b_t = \frac{2N}{N+1}\hat{x}_{t+1} - \frac{N+3}{N+1}\hat{\hat{x}}_{t+1}. \tag{2.1.15}$$

式 (2.1.15) 即为 a_t 的估计.

2. 二次指数平滑预测模型和参数估计

若时间序列$\{x_t, t=1,2,\cdots,n\}$具有线性趋势, 它还可以用二次指数平滑预测法进行预测, 所谓二次指数平滑预测法就是对一次指数平滑预测值序列再进行一次指数平滑预测得到的值. 令

$$\hat{x}_{n+1} = \alpha\, x_n + (1-\alpha)\hat{x}_n, \tag{2.1.16}$$

称式 (2.1.16) 是一次指数平滑值模型. 若对其再作一次指数平滑, 就得到二次指数平滑计算式

$$\hat{\hat{x}}_{n+1} = \alpha\, \hat{x}_n + (1-\alpha)\hat{\hat{x}}_n, \tag{2.1.17}$$

其中 $\hat{\hat{x}}_{n+1}$ 为二次指数平滑预测值, α 为平滑系数. 二次指数平滑的初始值也可取为初始时刻时间序列的实际值, 即 $\hat{\hat{x}}_0 = x_1$.

设时间序列可以向两侧无限延伸, 即$\{x_t, t=\cdots,-2,-1,0,1,2,\cdots,n\}$, 由式 (2.1.6) 知

$$\hat{x}_{n+1} = \alpha\sum_{i=0}^{\infty}(1-\alpha)^i x_{n-i}, \tag{2.1.18}$$

同理, 由式 (2.1.17) 有

$$\hat{\hat{x}}_{n+1} = \alpha\sum_{i=0}^{\infty}(1-\alpha)^i \hat{x}_{n-i}, \tag{2.1.19}$$

假设两侧无限延伸的时间序列具有线性趋势 $x_{t+T} = a_t + b_t T$, 由式 (2.1.18) 得

$$\begin{aligned}
\hat{x}_{n+1} &= \alpha\sum_{i=0}^{\infty}(1-\alpha)^i x_{n-i} = \alpha\sum_{i=0}^{\infty}(1-\alpha)^i(a_n - ib_n)\\
&= \alpha\sum_{i=0}^{\infty}(1-\alpha)^i a_n - \alpha\sum_{i=0}^{\infty}(1-\alpha)^i ib_n\\
&= a_n - \alpha(1-\alpha)b_n\sum_{i=1}^{\infty}(1-\alpha)^{i-1}i = a_n - \alpha(1-\alpha)b_n\left[\mathrm{d}\left(\sum_{i=1}^{\infty}t^i\right)/\mathrm{d}t\right]\Bigg|_{t=1-\partial}\\
&= a_n - \alpha(1-\alpha)b_n\left[\mathrm{d}\left(\frac{t}{1-t}\right)\Big/\mathrm{d}t\right]\Bigg|_{t=1-\alpha} = a_n - \frac{1-\alpha}{\alpha}b_n.
\end{aligned} \tag{2.1.20}$$

把式 (2.1.18) 代入式 (2.1.19) 得

$$\begin{aligned}
\hat{\hat{x}}_{n+1} &= \alpha\sum_{i=0}^{\infty}(1-\alpha)^i\left[\alpha\sum_{k=0}^{\infty}(1-\alpha)^k(a_n-(i+k)b_n)\right]\\
&= \alpha\sum_{i=0}^{\infty}(1-\alpha)^i\left[a_n - ib_n - \frac{1-\alpha}{\alpha}b_n\right] = a_n - \frac{2(1-\alpha)}{\alpha}b_n.
\end{aligned} \tag{2.1.21}$$

由式 (2.1.20) 和式 (2.1.21) 联合解方程组得

$$\begin{cases} a_n = 2\hat{x}_{n+1} - \hat{\hat{x}}_{n+1}, \\ b_n = \dfrac{1-\alpha}{\alpha}\left(\hat{x}_{n+1} - \hat{\hat{x}}_{n+1}\right). \end{cases} \tag{2.1.22}$$

式 (2.1.22) 即为 a_n, b_n 的估计.

2.2 回归分析预测模型

回归分析是英国生物学家 Galton 在研究人类身高的遗传特性时首先提出来的. 通过大量的统计资料, Galton 发现, 身材高的父亲, 他的儿子身材比其略矮; 反之, 身材矮的父亲, 他的儿子身材比其略高. 即人类的身高有向平均数靠近的倾向, 这种现象称之为回归. 回归分析预测方法是一种因果分析预测. 它是研究某一个随机变量与一个或几个变量之间的数量关系, 用一个或几个非随机变量来预测某一个随机变量的方法.

回归分析预测方法按回归方程所含的变量的多少分为一元回归和多元回归. 按回归方程的性质分为线性回归和非线性回归. 下面介绍多元线性回归预测.

设预测对象 y 受 m 个因素 $x_1, x_2, \cdots, x_m$ 的影响, y 与因素 $x_1, x_2, \cdots, x_m$ 之间满足以下线性关系

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m + \varepsilon, \tag{2.2.1}$$

若获得 n 次观察值 $(y_i, x_{i1}, x_{i2}, \cdots, x_{im})$, $i = 1, 2, \cdots, n$, 则有下式成立

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \cdots + \alpha_m x_{im} + \varepsilon_i, \qquad i = 1, 2, \cdots, n, \tag{2.2.2}$$

令

$$\boldsymbol{Y} = (y_1, y_2, \cdots, y_n)^{\mathrm{T}}, \qquad \boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)^{\mathrm{T}}, \qquad \boldsymbol{\alpha} = (\alpha_0, \alpha_1, \cdots, \alpha_m)^{\mathrm{T}},$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix},$$

则式 (2.2.2) 可写成

$$\boldsymbol{Y} = \boldsymbol{X\alpha} + \boldsymbol{\varepsilon}, \tag{2.2.3}$$

其中 $\boldsymbol{\alpha}$ 为未知参数, $\boldsymbol{Y}$ 为观察值向量, $\boldsymbol{X}$ 为观察值矩阵, $\boldsymbol{\varepsilon}$ 为随机误差向量. 一般假设

$$\operatorname{rank}(\boldsymbol{X}) = m+1, \qquad E(\boldsymbol{\varepsilon}) = \mathbf{0}, \qquad \operatorname{cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{I}_n, \tag{2.2.4}$$

其中 σ^2 为随机误差的方差, $\boldsymbol{I}_n$ 为 n 阶单位阵.

通常称满足式 (2.2.3)、(2.2.4) 的 $\boldsymbol{Y}$ 服从线性回归模型.

2.2.1 未知参数向量的最小二乘 (LS) 估计和性质

定义 2.2.1 对于线性模型 (2.2.3), 若存在 $\boldsymbol{Y}$ 的线性函数 $\hat{\boldsymbol{\alpha}}$, 使得

$$(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}})^{\mathrm{T}} (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}) = \min_{\alpha} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})^{\mathrm{T}} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha}),$$

则称 $\hat{\boldsymbol{\alpha}}$ 为未知参数向量 $\boldsymbol{\alpha}$ 的最小二乘估计.

定理 2.2.1 对于线性回归模型 (2.2.3), $\hat{\boldsymbol{\alpha}} = \left(\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$ 为未知参数向量 $\boldsymbol{\alpha}$ 的最小二乘估计.

证明 令 $Q = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})^{\mathrm{T}} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})$, Q 表示误差平方和. 即

$$Q = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y} - 2\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{X}^{\mathrm{T}} \boldsymbol{X}\boldsymbol{\alpha},$$

要使 Q 达到最小, 由极值存在的一阶条件知

$$\frac{\partial Q}{\partial \boldsymbol{\alpha}} = -2\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + 2\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X}\boldsymbol{\alpha} = \mathbf{0}, \tag{2.2.5}$$

因为 $\operatorname{rank}(\boldsymbol{X})=m+1$, 所以 $\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X}$ 的逆存在, 从而方程组 (2.2.5) 的解存在且唯一, 其解为

$$\hat{\boldsymbol{\alpha}} = \left(\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}.$$

下证 $\hat{\boldsymbol{\alpha}}$ 为 Q 的最优解.

设 $\tilde{\boldsymbol{\alpha}}$ 为 $\boldsymbol{\alpha}$ 的任意的其他估计, $\tilde{\boldsymbol{\alpha}}$ 不等于 $\hat{\boldsymbol{\alpha}}$, 则

$$\begin{aligned}(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}}) &= [(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}) + (\boldsymbol{X}\hat{\alpha} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}})]^{\mathrm{T}}[(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}) \\ &\quad + (\boldsymbol{X}\hat{\boldsymbol{\alpha}} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}})] \\ &= (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}) + (\boldsymbol{X}\hat{\boldsymbol{\alpha}} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}})^{\mathrm{T}}(\boldsymbol{X}\hat{\boldsymbol{\alpha}} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}}) \\ &\quad + 2(\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}})^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}).\end{aligned}$$

由式 (2.2.5) 得 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}} \boldsymbol{X}\hat{\boldsymbol{\alpha}}$, 所以上式第三项有

$$(\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}})^{\mathrm{T}} \boldsymbol{X}^{\mathrm{T}} (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}) = (\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}})^{\mathrm{T}} \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} - \boldsymbol{X}^{\mathrm{T}} \boldsymbol{X}\hat{\boldsymbol{\alpha}}\right) = 0,$$

且注意到 $\hat{\boldsymbol{\alpha}} \neq \tilde{\boldsymbol{\alpha}}$, 从而

$$(\boldsymbol{X}\hat{\boldsymbol{\alpha}} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}})^{\mathrm{T}}(\boldsymbol{X}\hat{\boldsymbol{\alpha}} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}}) > 0.$$

所以

$$(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\alpha}}) > (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}).$$

即

$$(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}) = \min_{\alpha}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha}).$$

所以 $\hat{\alpha}$ 为线性模型 (2.2.3) 的 LS 估计.

性质 2.2.1 $\hat{\boldsymbol{\alpha}}$ 为 $\boldsymbol{\alpha}$ 的无偏估计, 即 $\boldsymbol{E}(\hat{\boldsymbol{\alpha}}) = \boldsymbol{\alpha}$.

证明 因为 $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$, 且 $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$, 所以 $E\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\alpha}$. 则有

$$\begin{aligned}[E(\hat{\boldsymbol{\alpha}}) &= E[(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}] = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}E[\boldsymbol{Y}] \\ &= (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{\alpha} = \boldsymbol{\alpha}.\end{aligned}$$

性质 2.2.2 $\operatorname{cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}}) = \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$.

证明 因为 $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$, 且 $\operatorname{cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{I}_n$, 所以 $\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Y}) = \sigma^2\boldsymbol{I}_n$.

$$\begin{aligned}\operatorname{cov}(\hat{\alpha}, \hat{\alpha}) &= \operatorname{cov}((\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}, (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}) \\ &= (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Y})\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} \\ &= (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}} \cdot \sigma^2\boldsymbol{I}_n \cdot \boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} = \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}.\end{aligned}$$

性质 2.2.3 $\hat{\boldsymbol{\alpha}}$ 为 $\boldsymbol{\alpha}$ 的最小方差线性无偏估计, 即 $\operatorname{cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}}) \leqslant \operatorname{cov}(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}})$, 其中 $\tilde{\boldsymbol{\alpha}}$ 为 $\boldsymbol{\alpha}$ 的另一线性无偏估计.

证明 设 $\tilde{\boldsymbol{\alpha}} = \boldsymbol{A}\boldsymbol{Y}$ 为 $\boldsymbol{\alpha}$ 的另一线性无偏估计, 其中 $\boldsymbol{A}$ 为 $(m+1) \times n$ 矩阵. 即

$$E(\tilde{\boldsymbol{\alpha}}) = E(\boldsymbol{A}\boldsymbol{Y}) = E(\boldsymbol{A}(\boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon})) = \boldsymbol{\alpha},$$

所以有 $\boldsymbol{A}\boldsymbol{X}\boldsymbol{\alpha} = \boldsymbol{\alpha}$, 对任意的 $\boldsymbol{\alpha}$ 成立, 从而有 $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{I}_{m+1}$, 其中 $\boldsymbol{I}_{m+1}$ 表示 $m+1$ 阶单位阵. 注意到

$$\begin{aligned}\operatorname{cov}(\boldsymbol{A}\boldsymbol{Y}, \boldsymbol{A}\boldsymbol{Y}) &= E(\boldsymbol{A}\boldsymbol{Y} - E(\boldsymbol{A}\boldsymbol{Y}))(\boldsymbol{A}\boldsymbol{Y} - E(\boldsymbol{A}\boldsymbol{Y}))^{\mathrm{T}} \\ &= \boldsymbol{A}E(\boldsymbol{Y} - E(\boldsymbol{Y}))(\boldsymbol{Y} - E(\boldsymbol{Y}))^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A} \cdot \sigma^2\boldsymbol{I}_n \cdot \boldsymbol{A}^{\mathrm{T}} = \sigma^2\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}.\end{aligned}$$

由性质 2.2.2 知 $\operatorname{cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}}) = \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$, 所以

$$\begin{aligned}\operatorname{cov}(\boldsymbol{A}\boldsymbol{Y}, \boldsymbol{A}\boldsymbol{Y}) - \operatorname{cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}}) &= \sigma^2\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} - \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} \\ &= \sigma^2(\boldsymbol{A} - (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}) \\ &\quad \times (\boldsymbol{A} - (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}})^{\mathrm{T}} \geqslant 0.\end{aligned}$$

即 $\operatorname{cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}}) \leqslant \operatorname{cov}(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}})$.

2.2.2 随机误差方差的估计

因为 $\hat{\boldsymbol{\alpha}}$ 为 $\boldsymbol{\alpha}$ 的最小二乘估计, 所以 $\hat{\boldsymbol{Y}}=\boldsymbol{X}\hat{\boldsymbol{\alpha}}$ 为 $\boldsymbol{Y}$ 的估计值. 令 $\boldsymbol{e}=\boldsymbol{Y}-\hat{\boldsymbol{Y}}$, 则称 $\boldsymbol{e}$ 为残差向量. 显然有

$$\boldsymbol{e}=\boldsymbol{Y}-\hat{\boldsymbol{Y}}=\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\alpha}}=(\boldsymbol{I}_n-\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}})\boldsymbol{Y}=(\boldsymbol{I}_n-\boldsymbol{P})\boldsymbol{Y},$$

其中 $\boldsymbol{P}_{\boldsymbol{X}}=\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$, 显然 $\boldsymbol{P}_{\boldsymbol{X}}$ 为对称、幂等阵, 即 $\boldsymbol{P}_{\boldsymbol{X}}^{\mathrm{T}}=\boldsymbol{P}_{\boldsymbol{X}}, \boldsymbol{P}_{\boldsymbol{X}}^2=\boldsymbol{P}_{\boldsymbol{X}}$.

性质 2.2.4 残差向量 $\boldsymbol{e}$ 具有下列性质:

(1) $E(\boldsymbol{e})=\boldsymbol{0}$. (2.2.6)

(2) $\operatorname{cov}(\boldsymbol{e},\boldsymbol{e})=\sigma^2(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})$. (2.2.7)

(3) 若进一步假设 ε 服从多元正态分布, 即 $\varepsilon\sim N(\boldsymbol{0},\sigma^2\boldsymbol{I}_n)$, 则

$$\boldsymbol{e}\sim N(\boldsymbol{0},\ \sigma^2(\boldsymbol{I}_n-\boldsymbol{R}_{\boldsymbol{X}}))$$

证明 (1)

$$\begin{aligned}E(\boldsymbol{e})&=(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})E\boldsymbol{Y}=(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{X}\boldsymbol{\alpha}\\&=(\boldsymbol{I}_n-\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}})\boldsymbol{X}\boldsymbol{\alpha}=\boldsymbol{0}.\end{aligned}$$

(2)

$$\begin{aligned}\operatorname{cov}(\boldsymbol{e},\boldsymbol{e})&=\operatorname{cov}((\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{Y},(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{Y})\\&=(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Y})(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})^{\mathrm{T}}\\&=(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})\cdot\sigma^2\boldsymbol{I}_n\cdot(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}})^{\mathrm{T}}=\sigma^2(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}}).\end{aligned}$$

(3) 因为 $\varepsilon\sim N(\boldsymbol{0},\ \sigma^2\boldsymbol{I}_n)$, 所以 $\boldsymbol{Y}\sim N(\boldsymbol{X}\boldsymbol{\alpha},\ \sigma^2\boldsymbol{I}_n)$, 从而 $\sim N(0,\ \sigma^2(\boldsymbol{I}_n-\boldsymbol{P}_{\boldsymbol{X}}))$.

利用残差向量 $\boldsymbol{e}$ 可构造误差方差 σ^2 的估计. 先引进一个引理.

引理 2.2.1 设 $\boldsymbol{\xi}$ 为 n 维随机列向量, $E(\boldsymbol{\xi})=\boldsymbol{\mu}$, $\operatorname{cov}(\boldsymbol{\xi},\ \boldsymbol{\xi})=\boldsymbol{\Sigma}$, $\boldsymbol{A}$ 为 n 阶对称矩阵, 则有

$$E(\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi})=\operatorname{tr}(\boldsymbol{A}\boldsymbol{\Sigma})+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu},\tag{2.2.8}$$

其中 $\operatorname{tr}(\cdot)$ 表示矩阵的迹.

证明

$$\begin{aligned}E(\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi})&=E((\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{\xi}-\boldsymbol{\mu})+\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu}+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi}-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu})\\&=E((\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{\xi}-\boldsymbol{\mu}))+2E(\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu})-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu}\\&=E(\operatorname{tr}(\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{\xi}-\boldsymbol{\mu}))+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu}=E(\operatorname{tr}\boldsymbol{A}(\boldsymbol{\xi}-\boldsymbol{\mu})(\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}})+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu}\\&=\operatorname{tr}(\boldsymbol{A}E(\boldsymbol{\xi}-\boldsymbol{\mu})(\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}})+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu}=\operatorname{tr}(\boldsymbol{A}\boldsymbol{\Sigma})+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu}.\end{aligned}$$

定理 2.2.2 对于线性模型 (2.2.3), $\hat{\sigma}^2 = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}/(n-m-1)$ 为 σ^2 的一个无偏估计.

证明 由式 (2.2.5) 并注意到 $\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}}$ 为对称、幂等阵知

$$\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e} = \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})^{\mathrm{T}} \cdot (\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{Y} = \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{Y},$$

由引理 2.2.1 知 $E(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}) = (\boldsymbol{X}\alpha)^{\mathrm{T}}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})(\boldsymbol{X}\alpha) + \mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})\mathrm{cov}(\boldsymbol{Y},\ \boldsymbol{Y})$.

因为 $(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{X} = \boldsymbol{X} - \boldsymbol{X} = \boldsymbol{0}$, $\mathrm{cov}(\boldsymbol{Y},\ \boldsymbol{Y}) = \sigma^2\boldsymbol{I}_n$, 且

$$\begin{aligned}\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}}) &= \mathrm{tr}(\boldsymbol{I}_n) - \mathrm{tr}(\boldsymbol{P}_{\boldsymbol{X}}) = n - \mathrm{tr}\big(\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\big) \\ &= n - \mathrm{tr}\left((\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right) \\ &= n - \mathrm{tr}(\boldsymbol{I}_{m+1}) = n - m - 1.\end{aligned}$$

所以 $E(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}) = (n-m-1)\sigma^2$, 从而 $E\hat{\sigma}^2 = E\left(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}/(n-m-1)\right) = \sigma^2$.

引理 2.2.2[49] 设 $\boldsymbol{\xi}$ 为 n 维正态随机列向量, $\boldsymbol{\xi} \sim N(\boldsymbol{\mu},\ \boldsymbol{I}_n)$, $\boldsymbol{A}$ 为 n 阶对称矩阵, 则有 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi} \sim \chi^2(r,\ \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\mu})$ 的充要条件是 $\boldsymbol{A}$ 为幂等矩阵, $\mathrm{rank}(\boldsymbol{A}) = r$.

引理 2.2.3[49] 设 $\boldsymbol{\xi}$ 为 n 维正态随机列向量, $\boldsymbol{\xi} \sim N(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ 为 n 阶正定矩阵, $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶对称矩阵, $\boldsymbol{C}$ 为 $m \times n$ 的矩阵, 则有

(1) 若 $\boldsymbol{A\Sigma B} = \boldsymbol{0}$, 则 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi}$ 与 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\xi}$ 相互独立.

(2) 若 $\boldsymbol{C\Sigma A} = \boldsymbol{0}$, 则 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi}$ 与 $\boldsymbol{C}\boldsymbol{\xi}$ 相互独立.

定理 2.2.3 对于线性模型 (2.2.3), 若 $\boldsymbol{\varepsilon} \sim N\left(\boldsymbol{0}, \sigma^2\boldsymbol{I}_n\right)$, 则有

(1) $\dfrac{(n-m-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-m-1)$. (2.2.9)

(2) $\hat{\boldsymbol{\alpha}}$ 与 $\hat{\sigma}^2$ 相互独立.

证明 (1) 令 $\boldsymbol{Z} = \dfrac{\boldsymbol{Y} - \boldsymbol{X\alpha}}{\sigma}$,

$$\begin{aligned}\frac{(n-m-1)\hat{\sigma}^2}{\sigma^2} = \frac{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}}{\sigma^2} &= \frac{\boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{Y}}{\sigma^2} \\ &= \frac{(\boldsymbol{Y} - \boldsymbol{X\alpha})^{\mathrm{T}}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})(\boldsymbol{Y} - \boldsymbol{X\alpha})}{\sigma^2} \\ &= \boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{Z}.\end{aligned}$$

因为 $\boldsymbol{\varepsilon} \sim N\left(\boldsymbol{0},\ \sigma^2\boldsymbol{I}_n\right)$, 所以 $\boldsymbol{Z} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I}_n)$. 由 $\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}}$ 为对称幂等阵知

$$\mathrm{rank}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}}) = \mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}}) = n - m - 1,$$

由引理 2.2.2 知 $\boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{Z} \sim \chi^2(n-m-1)$, 即结论成立.

(2) 因为 $\boldsymbol{\varepsilon} \sim N\left(\mathbf{0},\ \sigma^2\boldsymbol{I}_n\right)$, 所以 $\boldsymbol{Y} \sim N\left(\boldsymbol{X\alpha},\ \sigma^2\boldsymbol{I}_n\right)$. 而 $\hat{\boldsymbol{\alpha}} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$ 为 $\boldsymbol{Y}$ 的线性型,

$$\hat{\sigma}^2 = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}/(n-m-1) = \boldsymbol{Y}^{\mathrm{T}}\left[(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})/(n-m-1)\right]\boldsymbol{Y}$$

为 $\boldsymbol{Y}$ 的二次型, 且

$$\begin{aligned}
&(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\cdot\sigma^2\boldsymbol{I}_n\cdot\left[(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}})/(n-m-1)\right]\\
&\quad = \sigma^2\cdot\left[(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}} - (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\cdot\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\right]\\
&\quad /(n-m-1) = \mathbf{0}.
\end{aligned}$$

由引理 2.2.3 知 $\hat{\boldsymbol{\alpha}}$ 与 $\hat{\sigma}^2$ 相互独立.

2.2.3　回归预测模型的统计假设检验

1.　F 检验

记

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,\quad \mathrm{SST} = \sum_{i=1}^{n}(y_i - \bar{y})^2,$$

$$\mathrm{SSR} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2,\quad \mathrm{SSE} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2.$$

则称 SST 为总离差平方和, SSR 为回归平方和, SSE 为残差平方和. 其中 $\hat{y}_i = \hat{\alpha}_0 + \sum_{j=1}^{m}\hat{\alpha}_i x_{ij}$ 为 y_i 的估计.

下面考察总离差平方和的分解式

$$\begin{aligned}
\sum_{i=1}^{n}(y_i - \bar{y})^2 &= \sum_{i=1}^{n}\left[(y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})\right]^2\\
&= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + 2\sum_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \bar{y}).
\end{aligned}$$

上式的第三项为零. 实际上, 令 $Q = (\boldsymbol{Y} - \hat{\boldsymbol{Y}})^{\mathrm{T}}(\boldsymbol{Y} - \hat{\boldsymbol{Y}})$ 为残差平方和, 由正规方程组 $\dfrac{\partial Q}{\partial \alpha_0} = 0$ 得

$$(\boldsymbol{Y} - \hat{\boldsymbol{Y}})^{\mathrm{T}}\boldsymbol{R} = 0,$$

其中 $\boldsymbol{R} = (1\ 1\ \cdots\ 1)^{\mathrm{T}}$, 并注意到估计式 $\hat{\boldsymbol{\alpha}} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$, 所以有

$$\begin{aligned}\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\bar{y}) &= (\boldsymbol{Y}-\hat{\boldsymbol{Y}})^{\mathrm{T}}(\hat{\boldsymbol{Y}}-\bar{y}\boldsymbol{R})=(\boldsymbol{Y}-\hat{\boldsymbol{Y}})^{\mathrm{T}}\hat{\boldsymbol{Y}}-\bar{y}(\boldsymbol{Y}-\hat{\boldsymbol{Y}})^{\mathrm{T}}\boldsymbol{R}\\ &= (\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})\hat{\boldsymbol{\alpha}}-\bar{y}(\boldsymbol{Y}-\hat{\boldsymbol{Y}})^{\mathrm{T}}\boldsymbol{R}\\ &= (\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})\hat{\boldsymbol{\alpha}}-\bar{y}(\boldsymbol{Y}-\hat{\boldsymbol{Y}})^{\mathrm{T}}\boldsymbol{R}\\ &= (\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}-\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})\hat{\boldsymbol{\alpha}}-\bar{y}(\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{R}-\hat{\boldsymbol{Y}}^{\mathrm{T}}\boldsymbol{R})\\ &= 0.\end{aligned}$$

则有

$$\mathrm{SST}=\mathrm{SSR}+\mathrm{SSE}. \tag{2.2.10}$$

而 $\mathrm{SSE}=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}$, 由定理 2.2.3 知 $\dfrac{\mathrm{SSE}}{\sigma^2}\sim\chi^2(n-m-1)$.

在假设 $H_0:\ \boldsymbol{\alpha}=\boldsymbol{0}$ 的条件下, 同理可证 $\dfrac{\mathrm{SST}}{\sigma^2}\sim\chi^2(n-1)$, $\dfrac{\mathrm{SSR}}{\sigma^2}\sim\chi^2(m)$. 于是构造 F 统计量

$$F=\frac{\mathrm{SSR}/m}{\mathrm{SSE}/(n-m-1)}\sim F(m,n-m-1).$$

对给定的显著性水平 β, 由 F 分布表可查得临界值 F_β, 根据给定的样本观察值计算 F 值, 若 $F\leqslant F_\beta$, 则接受假设 $H_0:\ \boldsymbol{\alpha}=\boldsymbol{0}$. 这表明预测对象 y 与 m 个影响因素 $x_1,x_2,\cdots,x_m$ 的线性关系不显著. 反之, 若 $F>F_\beta$, 则拒绝假设 H_0. 这表明 y 与影响因素 $x_1,x_2,\cdots,x_m$ 的线性关系显著.

2. *t* 检验

记 $\boldsymbol{C}=\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}=(c_{ij})_{(m+1)\times(m+1)}$, 在 $\varepsilon\sim N\left(0,\ \sigma^2\boldsymbol{I}_n\right)$ 和假设 $H_0:\ \alpha_i=0$ 的条件下, 由性质 2.2.1、性质 2.2.2、定理 2.2.3 知

$$\hat{\alpha}_i\sim N\left(0,\sigma^2c_{ii}\right),\qquad \hat{\alpha}_i\text{与}\hat{\sigma}^2\text{相互独立}.$$

所以构造 t 统计量

$$T_i=\frac{\hat{\alpha}_i}{\sqrt{c_{ii}}\sqrt{\mathrm{SSE}/(n-m-1)}}\sim t\left(n-m-1\right),\qquad i=0,1,2,\cdots,m.$$

对给定的显著性水平 β, 由 t 分布表可查得临界值 t_β, 根据给定的样本观察值计算 T_i 值.

若 $|T_i|\leqslant t_\beta$, 则接受假设 $H_0:\ \alpha_i=0$. 这表明预测对象 y 与第 i 个影响因素 x_i 的线性关系不显著. 此时需要把 x_i 从回归模型中剔除, 重新建立回归模型.

反之, 若 $|T_i|>t_\beta$, 则拒绝假设 H_0. 这表明 y 与影响因素 x_i 的线性关系显著.

3. $D-W$ 检验

在建立回归模型时, 假设随机误差项之间互不相关, 即 $\operatorname{cov}(\varepsilon_i,\varepsilon_j)=0$, $i\neq j$, 若该假设不成立, 将会使线性回归模型的最小乘估计不再具有最小方差性, 从而使预测失效. 因此, 有必要对回归模型进行序列相关检验, D-W 检验步骤如下:

第一步　计算 D-W 检验统计量

$$DW=\sum_{i=2}^{n}(e_i-e_{i-1})^2\bigg/\sum_{i=2}^{n}e_i^2,$$

其中 $e_i=y_i-\hat{y}_i$, $i=1, 2, \cdots, n$.

第二步　查 D-W 检验表. 给定的显著性水平 β, 根据样本个数 n 和自变量个数 m 可得临界值 d_l, d_u.

第三步　比较判别.

当 $DW\in(d_l,\ d_u)$ 或 $DW\in(4-d_u,\ 4-d_l)$ 时, 检验无结论, 此时或增加样本个数或增加自变量个数;

当 $DW\in(4-d_l,\ \ 4)$ 时, 表明随机误差项之间出现负相关;

当 $DW\in(0,\ \ d_l)$ 时, 表明随机误差项之间出现正相关;

当 $DW\in(d_u,\ 4-d_u)$ 时, 表明随机误差项之间互不相关.

4. 多重共线性检验

在建立回归模型时, 假设 m 个影响因素 $x_1,x_2,\cdots,x_m$ 之间相互独立, 即认为 $\operatorname{rank}(\boldsymbol{X})=m+1$. 若影响因素之间存在完全的共线性时, $\operatorname{rank}(\boldsymbol{X})<m+1$, 将会使回归模型参数的最小乘估计不能计算出来. 在实际预测过程中, 若 $\left|\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right|\approx 0$, 表明影响因素之间多重共线性现象. 一般可采用如下条件数判别法:

设 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 的 $m+1$ 个特征根为 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{m+1}$, 则称 $k=\lambda_1/\lambda_{m+1}$ 为方阵 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 的条件数.

一般认为, 若 $k\in(0,100)$, 影响因素之间不存在共线性, 若 $k\in(100,1000)$, 则影响因素之间存在较强的共线性, 若 $k\in(1000,+\infty)$, 则影响因素之间存在严重的共线性.

2.2.4 回归模型预测方法

若建立的回归模型通过各种检验后, 则可以利用估计出的参数进行预测. 回归预测分为点预测和区间预测.

1. 点预测

若给定 m 个因素的一组值, 设 $x_1=x_1^{(0)},x_2=x_2^{(0)},\cdots,x_m=x_m^{(0)}$, 则预测对象

y 的点预测值为

$$\hat{y}_0 = \hat{\alpha}_0 + \sum_{i=1}^{m} \hat{\alpha}_i x_i^{(0)}. \tag{2.2.11}$$

2. 区间预测

令 $\boldsymbol{X}_0 = \left(1,\ x_1^{(0)}, x_2^{(0)}, \cdots, x_m^{(0)}\right)^{\mathrm{T}}$, $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0, \hat{\alpha}_1, \cdots, \hat{\alpha}_m)^{\mathrm{T}}$, 则 (2.2.11) 可表为 $\hat{y}_0 = \boldsymbol{X}_0^{\mathrm{T}}\hat{\boldsymbol{\alpha}}$, 在 ε 服从多元正态分布假设条件下, 则有

$$\hat{y}_0 \sim N\left(\boldsymbol{X}_0^{\mathrm{T}}\boldsymbol{\alpha},\ \sigma^2 \boldsymbol{X}_0^{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0\right), \qquad y_0 \sim N\left(\boldsymbol{X}_0^{\mathrm{T}}\boldsymbol{\alpha}, \sigma^2\right).$$

y_0 是没有观察到的总体的特定值, 可以认为 $\hat{y}_0$ 和 y_0 相互独立, 所以

$$y_0 - \hat{y}_0 \sim N\ \left(0,\ \ \sigma^2\left[1 + \boldsymbol{X}_0^{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0\right]\ \right).$$

即

$$(y_0 - \hat{y}_0)\Big/\sqrt{\sigma^2\left[1 + \boldsymbol{X}_0^{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0\right]} \sim N(\ 0,\ \ 1\ ).$$

由定理 2.2.3 知 $\dfrac{(n-m-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-m-1)$, 所以

$$\frac{(y_0 - \hat{y}_0)}{\hat{\sigma}\sqrt{1 + \boldsymbol{X}_0^{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0}} \sim t(n-m-1).$$

从而 y_0 的 $1-\beta$ 的置信区间预测为

$$\hat{y}_0 - t_{\beta/2}\hat{\sigma}\sqrt{1 + \boldsymbol{X}_0^{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0} \leqslant y_0 \leqslant \hat{y}_0 + t_{\beta/2}\hat{\sigma}\sqrt{1 + \boldsymbol{X}_0^{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0}. \tag{2.2.12}$$

2.3 随机时间序列预测模型

美国著名的统计学家 Box 和英国著名的统计学家 Jenkins 于 1968 年提出随机时间序列模型的识别、参数估计和诊断检验的建模方法, 并出版了专著《时间序列分析 —— 预测和控制》[50], 该书对随机时间序列分析的理论分析和应用作了系统论述, 特别是计算机技术的迅速发展, 为随机时间序列分析的建模和预测提供了强有力的工具.

2.3.1 平稳时间序列

对固定的时间 t, y_t 为一个随机变量, 则称 $\{y_t, t = 1, 2, \cdots\}$ 为一个随机变量序列. 称 $\{y_t, t \in (-\infty, +\infty)\}$ 为一随机过程.

定义 2.3.1 $\{y_t, t=1,2,\cdots\}$ 为一个随机时间序列, 记 $\mu_t = Ey_t$, 则称 μ_t 为 $\{y_t, t=1,2,\cdots\}$ 的均值函数.

定义 2.3.2 $\{y_t, t=1,2,\cdots\}$ 为一个随机时间序列, 记 $\gamma_{s,\,t} = E(y_s-\mu_s)(y_t-\mu_t)$, 则称 $\gamma_{s,\,t}$ 为 $\{y_t, t=1,2,\cdots\}$ 的自协方差函数. 特别地, $\gamma_{t,\,t}$ 为 $\{y_t, t=1,2,\cdots\}$ 的方差函数.

定义 2.3.3 记 $\rho_{s,\,t} = \gamma_{s,\,t}/\sqrt{\gamma_{s,\,s}\cdot\gamma_{t,\,t}}$, 则称 $\rho_{s,\,t}$ 为 $\{y_t, t=1,2,\cdots\}$ 的自相关函数.

定义 2.3.4 $\{y_t, t=1,2,\cdots\}$ 为一个随机时间序列, 若满足 $Ey_t=\mu$, μ 为常数, $t=1,2,\cdots$, 且 $\gamma_{s,\,t}$ 只与时间差 s - t 有关, 记作 γ_s - t, 则称 $\{y_t, t=1,2,\cdots\}$ 为宽平稳随机时间序列.

宽平稳随机时间序列自协方差函数具有如下性质.

性质 2.3.1 $\gamma_0 = E(y_t-\mu)^2$ 为平稳随机时间序列的方差, 即任何时刻 t 的方差是相同的.

性质 2.3.2 $\gamma_{-k}=\gamma_k$, 自协方差函数为 k 的偶函数.

证明 由定义 2.3.2 知

$$\gamma_k = E(y_{t+k}-\mu)(y_t-\mu) = E(y_t-\mu)(y_{t+k}-\mu) = \gamma_{-k}.$$

性质 2.3.3 $|\gamma_k| \leqslant \gamma_0$, $k=\pm1, \pm2, \cdots$.

证明 对任意实数 λ,

$$\begin{aligned}
&E[\lambda(y_{t+k}-\mu)+(y_t-\mu)]^2\\
=&\lambda^2E[(y_{t+k}-\mu)^2+2\lambda E[(y_t-\mu)(y_{t+k}-\mu)+E[(y_t-\mu)^2\\
=&\lambda^2\gamma_0+2\lambda\gamma_k+\gamma_0\\
\geqslant& 0.
\end{aligned}$$

上式左边是关于实数 λ 的一元二次函数, 对任意实数 λ 均满足非负性, 且 $\gamma_0 \geqslant 0$, 所以有

$$(2\gamma_k)^2-4\gamma_0^2\leqslant 0,\quad 即\, |\gamma_k|\leqslant\gamma_0,\quad k=\pm1, \pm2, \cdots.$$

性质 2.3.4 对任意实数 $a_1, a_2, \cdots, a_n$, 有

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_ia_j\gamma_{i-j}\geqslant 0,$$

等号成立当且仅当 $a_1=a_2=\cdots=a_n=0$.

证明 $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_ia_j\gamma_{i-j} = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_ia_jE(y_i-\mu)(y_j-\mu) = E\left(\sum\limits_{i=1}^{n}a_i(y_i-\mu)\right)^2 \geqslant 0$, 显然等号成立当且仅当 $a_1=a_2=\cdots=a_n=0$.

宽平稳随机时间序列自相关函数具有如下性质：

性质 2.3.5 $\rho_0 = 1$.

证明 由定义 2.3.3 知 $\rho_0 = \gamma_0/\sqrt{\gamma_0 \cdot \gamma_0} = 1$.

性质 2.3.6 $\rho_{-k} = \rho_k$, 自相关函数为 k 的偶函数.

证明 由定义 2.3.3 知

$$\rho_k = \gamma_k/\gamma_0.$$

而 $\gamma_{-k} = \gamma_k$, 所以 $\rho_{-k} = \rho_k$.

性质 2.3.7 $|\rho_k| \leqslant 1$, $k = 0, \pm 1, \pm 2, \cdots$.

证明 因为 $\rho_k = \gamma_k/\gamma_0$ 且 $|\gamma_k| \leqslant \gamma_0$, $k = \pm 1, \pm 2, \cdots$, 所以

$$|\rho_k| \leqslant 1, \qquad k = 0, \pm 1, \pm 2, \cdots.$$

2.3.2 平稳随机时间序列模型及识别

1. 平稳随机时间序列模型的分类

定义 2.3.5 设 $\{\varepsilon_t, t = 1, 2, \cdots\}$ 为平稳随机时间序列, 且满足 $E\varepsilon_t = 0$, $t = 1, 2, \cdots$, 且 $\gamma_k = 0$, $k \neq 0$, $\gamma_0 = \sigma^2 > 0$, 则称 $\{\varepsilon_t, t = 1, 2, \cdots\}$ 为白噪声时间序列.

平稳随机时间序列模型可分为三类：

(1) p 阶自回归模型 AR(p)

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \tag{2.3.1}$$

其中 $\{\varepsilon_t, t = 1, 2, \cdots\}$ 为白噪声时间序列, 且满足 $E\varepsilon_t y_{t-i} = 0$, ϕ_1, $\phi_2, \cdots, \phi_p$ 为自回归系数.

(2) q 阶滑动平均模型 MA(q)

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \tag{2.3.2}$$

其中 $\{\varepsilon_t, t = 1, 2, \cdots\}$ 为白噪声时间序列. $\theta_1, \theta_2, \cdots, \theta_q$ 为滑动平均系数.

(3) 自回归 – 滑动平均模型 ARMA(p, q)

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \tag{2.3.3}$$

其中 $\{\varepsilon_t, t = 1, 2, \cdots\}$ 为白噪声时间序列, 且满足 $E\varepsilon_t y_{t-i} = 0$. ϕ_1, $\phi_2, \cdots, \phi_p$ 和 $\theta_1, \theta_2, \cdots, \theta_q$ 为 ARMA(p, q) 模型待估参数.

2. 平稳随机时间序列模型的识别

不失一般性, 在以下的讨论中均假定平稳随机时间序列 $Ey_t = 0$, $t = 1, 2, \cdots$, 因为若 $Ey_t \neq 0$, 即只要将它的均值减去即可转化为上述情形.

(1) p 阶自回归模型 AR(p) 的识别. AR(p) 模型为

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t.$$

则有

$$\begin{aligned}\gamma_k = Ey_{t+k}y_t &= E\left(\phi_1 y_{t+k-1} + \phi_2 y_{t+k-2} + \cdots + \phi_p y_{t+k-p} + \varepsilon_{t+k}\right) y_t \\ &= \phi_1 r_{k-1} + \phi_2 r_{k-2} + \cdots + \phi_p r_{k-p}.\end{aligned}$$

因为 $\rho_k = \gamma_k/\gamma_0$ 且 $\rho_{-k} = \rho_k$, 所以有

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}.$$

这是一个差分方程, 在 AR(p) 模型平稳域内被负指数函数控制, 称这种序列为拖尾序列. 因此自相关函数拖尾是 AR(p) 模型的一种特征.

AR(p) 的偏相关函数就是在给定 $y_{t-1}, y_{t-2}, \cdots, y_{t-k+1}$ 的条件下, y_t 与 y_{t-k} 之间的条件相关函数. 设 y_t 的预测方程是 $\hat{y}_t = \phi_{k1}y_{t-1} + \phi_{k2}y_{t-2} + \cdots + \phi_{kk}y_{t-k}$, 当 $k > p$ 时,

$$\begin{aligned}E(e_t^2) &= E[y_t - \hat{y}_t]^2 \\ &= E\left[\varepsilon_t + \sum_{i=1}^{p}(\phi_i - \phi_{ki})y_{t-i} - \sum_{i=p+1}^{k}\phi_{ki}y_{t-i}\right]^2 \\ &= \sigma_\varepsilon^2 + E\left[\sum_{i=1}^{p}(\phi_i - \phi_{ki})y_{t-i} - \sum_{i=p+1}^{k}\phi_{ki}y_{t-i}\right]^2 \\ &\quad + 2\left(\sum_{i=1}^{p}(\phi_i - \phi_{ki})E\varepsilon_t y_{t-i} - \sum_{i=p+1}^{k}\phi_i E\varepsilon_t y_{t-i}\right).\end{aligned}$$

因为当 $i > 0$ 时, $E\varepsilon_t y_{t-i}$=0, 所以

$$E(e_t^2) = \sigma_\varepsilon^2 + E\left[\sum_{i=1}^{p}(\phi_i - \phi_{ki})y_{t-i} - \sum_{i=p+1}^{k}\phi_{ki}y_{t-i}\right]^2.$$

要使 $E(e_t^2)$ 达到最小, 只有 $E\left[\sum\limits_{i=1}^{p}(\phi_i-\phi_{ki})y_{t-i}-\sum\limits_{i=p+1}^{k}\phi_{ki}y_{t-i}\right]^2=0$. 即

$$\phi_{ki}=\begin{cases}\phi_i, & i=1,\ 2,\ \cdots,\ p.\\ 0, & i=p+1,\cdots,k.\end{cases} \tag{2.3.4}$$

上式表明满足 AR(p) 模型的序列 $\{y_t,t=1,\ 2,\ \cdots\}$ 的偏相关函数 $\{\phi_{kk}\}$ 为 p 步截尾序列, 即当 $k>p$ 时, $\phi_{kk}=0$. 因此偏相关函数 p 步截尾是 AR(p) 模型的另一种特征.

(2) q 阶滑动平均模型 MA(q) 的识别. MA(q) 模型为

$$y_t=\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\cdots+\theta_q\varepsilon_{t-q},$$

其中 $\{\varepsilon_t,t=1,2,\cdots\}$ 为白噪声时间序列. 即

$$E(\varepsilon_t,\varepsilon_{t-k})=\begin{cases}\sigma^2, & k=0,\\ 0\ , & k\neq 0.\end{cases}$$

则有

$$\begin{aligned}\gamma_k=Ey_ty_{t-k}&=E\left[\left(\varepsilon_t+\sum_{i=1}^{q}\theta_iy_{t-i}\right)\left(\varepsilon_{t-k}+\sum_{i=1}^{q}\theta_iy_{t-k-i}\right)\right]\\&=\begin{cases}\sigma^2(1+\theta_1^2+\ \theta_2^2+\cdots+\theta_q^2), & k=0,\\ \sigma^2(\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q), & k\in[1,\ q],\\ 0, & k\in(q,\ +\infty).\end{cases}\end{aligned} \tag{2.3.5}$$

因为 $\rho_k=\gamma_k/\gamma_0$, 所以有

$$\rho_k=\begin{cases}1, & k=0,\\ \dfrac{\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q}{1+\theta_1^2+\ \theta_2^2+\cdots+\theta_q^2}, & k\in[1,\ q],\\ 0, & k\in(q,\ +\infty).\end{cases} \tag{2.3.6}$$

上式表明, 满足 MA(q) 模型的序列 $\{y_t,t=1,\ 2,\ \cdots\}$ 的自相关函数 $\{\rho_k\}$ 为 q 步截尾序列, 即当 $k>q$ 时, $\rho_k=0$. 因此, 自相关函数为 q 步截尾是 MA(q) 模型的一种特征.

对于 MA(q) 模型, 可以证明其偏相关函数为拖尾序列.

(3) 自回归 - 滑动平均模型 ARMA(p, q) 的识别. 根据上面的讨论, 可以得出如下识别准则:

若平稳随机时间序列 $\{y_t\}$ 的偏相关函数为截尾的, 自相关函数为拖尾的, 则 $\{y_t\}$ 为 AR 序列; 若平稳随机时间序列 $\{y_t\}$ 的自相关函数为截尾的, 偏相关函数为拖尾的, 则 $\{y_t\}$ 为 MA 序列; 若平稳随机时间序列 $\{y_t\}$ 的自相关函数和偏相关函数均为拖尾的, 则 $\{y_t\}$ 为 ARMA 序列.

2.3.3 平稳随机时间序列模型的参数估计

1. p 阶自回归模型 AR(p) 的参数估计

设 $\hat{y}_t = \phi_{k1}y_{t-1} + \phi_{k2}y_{t-2} + \cdots + \phi_{kk}y_{t-k}$ 是 y_t 的估计, ϕ_{k1}, $\phi_{k2}, \cdots, \phi_{kk}$ 是待估计参数, 通常采用误差最小的原则. 即

$$\min Q = E(e_t^2) = E\left[y_t - \sum_{i=1}^{k} \phi_{ki} y_{t-i}\right]^2 .$$

显然 Q 为 ϕ_{k1}, $\phi_{k2}, \cdots, \phi_{kk}$ 的函数, 要使 Q 达到最小, 由一阶条件 $\dfrac{\partial Q}{\partial \phi_{ki}} = 0$, $i = 1, 2, \cdots, k$ 得

$$\begin{cases} \phi_{k1} + \rho_1\phi_{k2} + \cdots + \rho_{k-1}\phi_{kk} = \rho_1 , \\ \rho_1\phi_{k1} + \phi_{k2} + \cdots + \rho_{k-2}\phi_{kk} = \rho_2, \\ \cdots\cdots\cdots\cdots \\ \rho_{k-1}\phi_{k1} + \rho_{k-2}\phi_{k2} + \cdots + \phi_{kk} = \rho_k. \end{cases} \tag{2.3.7}$$

此方程称为 Yule-Walker 方程.

定义 2.3.6 令

$$\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t, \qquad \hat{r}_k = \frac{1}{n}\sum_{t=1}^{n-k}(y_{t+k} - \bar{y})(y_t - \bar{y}), \qquad \hat{\rho}_k = \frac{\hat{r}_k}{\hat{r}_0}.$$

则称 $\bar{y}$ 为随机时间序列样本均值, $\hat{r}_k$ 为随机时间序列样本自协方差函数, $\hat{\rho}_k$ 为随机时间序列样本自相关函数.

只要以样本自相关函数 $\{\hat{\rho}_k\}$ 代替总体自相关函数 $\{\rho_k\}$, 就可以得到参数 ϕ_{k1}, $\phi_{k2}, \cdots, \phi_{kk}$ 的估计为

$$\begin{pmatrix} \hat{\phi}_{k1} \\ \hat{\phi}_{k2} \\ \vdots \\ \hat{\phi}_{kk} \end{pmatrix} = \begin{bmatrix} 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{k-1} \\ \hat{\rho}_1 & 1 & \cdots & \hat{\rho}_{k-2} \\ \vdots & \vdots & & \vdots \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \cdots & 1 \end{bmatrix}^{-1} \begin{pmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \vdots \\ \hat{\rho}_k \end{pmatrix}. \tag{2.3.8}$$

2. q 阶滑动平均模型 MA(q) 的参数估计

由式 (2.3.5) 就可以得到参数 $\theta_1,\theta_2,\cdots,\theta_q$ 的估计. 即当 $k\in[1,\quad q]$ 时, $\sigma^2(\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q)=\gamma_k$, 当 $k=0$ 时, $\sigma^2(1+\theta_1^2+\ \theta_2^2+\cdots+\theta_q^2)=\gamma_0$, 只要以样本自协方差函数 $\{\hat{\gamma}_k\}$ 代替总体自相关函数 $\{\gamma_k\}$ 得

$$\begin{cases} \theta_k=\dfrac{\hat{\gamma}_k}{\sigma^2}-(\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q)\,, & k=1,2,\cdots,\ q, \\ \sigma^2=\dfrac{\hat{\gamma}_0}{1+\theta_1^2+\ \theta_2^2+\cdots+\theta_q^2}. \end{cases} \tag{2.3.9}$$

上式是 q+1 个参数的非线性方程组, 通常用迭代法求解.

3. 自回归 - 滑动平均模型 ARMA(p,q) 的参数估计

在 ARMA(p,q) 模型中, 有 $\phi_1,\ \phi_2,\cdots,\phi_p,\ \theta_1,\theta_2,\cdots,\theta_q,\ \sigma^2$ 共 $p+q$+1 个参数需要估计. 下面介绍它们的计算步骤.

第一步 由扩充的 Yule-Walker 方程, 就可以得到参数 $\phi_1,\ \phi_2,\cdots,\phi_p$ 的估计为

$$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{pmatrix}=\begin{bmatrix} \hat{\rho}_q & \hat{\rho}_{q-1} & \cdots & \hat{\rho}_{q-p+1} \\ \hat{\rho}_{q+1} & \hat{\rho}_q & \cdots & \hat{\rho}_{q-p} \\ \vdots & \vdots & & \vdots \\ \hat{\rho}_{q+p} & \hat{\rho}_{q+p-1} & \cdots & \hat{\rho}_q \end{bmatrix}^{-1}\begin{pmatrix} \hat{\rho}_q \\ \hat{\rho}_{q+1} \\ \vdots \\ \hat{\rho}_{q+p} \end{pmatrix},$$

其中 $\{\hat{\rho}_k\}$ 为样本自相关函数.

第二步 令

$$\tilde{y}_t=y_t-\left(\hat{\phi}_1y_{t-1}+\hat{\phi}_2y_{t-2}+\cdots+\hat{\phi}_py_{t-p}\right),$$

则 $\{\tilde{y}_t\}$ 的样本自协方差函数 $\{\hat{\tilde{\gamma}}_k\}$ 可以由 $\{y_t\}$ 的样本自协方差函数 $\{\hat{\gamma}_k\}$ 来表达.

因为 $\tilde{\gamma}_k=E\tilde{y}_t\tilde{y}_{t-k}=\sum\limits_{i=0}^{p}\sum\limits_{j=0}^{p}\hat{\phi}_i\hat{\phi}_jEy_{t-k-j}y_{t-i}=\sum\limits_{i=0}^{p}\sum\limits_{j=0}^{p}\hat{\phi}_i\hat{\phi}_j\gamma_{k+j-i}$, 所以有

$$\hat{\tilde{\gamma}}_k=\sum_{i=0}^{p}\sum_{j=0}^{p}\hat{\phi}_i\hat{\phi}_j\hat{\gamma}_{k+j-i}.$$

第三步 将 ARMA 模型化为滑动平均模型 MA(q).

$$\tilde{y}_t=\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\cdots+\theta_q\varepsilon_{t-q},$$

再由滑动平均模型 MA(q) 的参数估计法可以获得 $\theta_1,\theta_2,\cdots,\theta_q,\ \sigma^2$ 共 q+1 个参数的估计.

2.3.4 平稳随机时间序列模型的预测方法

设 $\{y_t\}$ 为平稳随机时间序列, 实际的观察值为 $y_1, y_2, \cdots, y_t$, 现在利用 $y_1, y_2, \cdots, y_t$ 来预测 y_{t+l} 的值, 记为 $\hat{y}_t(l)$, 其中 l 表示预测的长度. 预测的准则就是要预测误差达到最小, 即

$$\min E\left(y_{t+l} - \hat{y}_t(l)\right)^2.$$

在上述准则下, 可以证明最佳预测 $\hat{y}_t(l)$ 是在 $y_1, y_2, \cdots, y_t$ 发生的条件下 y_{t+l} 的条件数学期望. 即

$$\hat{y}_t(l) = E\left(y_{t+l} \,|y_1, y_2, \cdots, y_t\right),$$

在预测中, 条件数学期望满足

$$E\left(y_{t+l} \,|y_1, y_2, \cdots, y_t\right) = \begin{cases} y_{t+l}, & l \leqslant 0, \\ \hat{y}_t(l), & l > 0, \end{cases}$$

$$E\left(\varepsilon_{t+l} \,|y_1, y_2, \cdots, y_t\right) = \begin{cases} \varepsilon_{t+l}\,, & l \leqslant 0, \\ 0\,, & l > 0. \end{cases}$$

因此有 ARMA(p,q) 模型最佳预测 $\hat{y}_t(l)$ 的通式为

$$\begin{aligned} \hat{y}_t(l) &= E\left(\sum_{i=1}^{p} \hat{\phi}_i y_{t+l-i} + \varepsilon_{t+l} + \sum_{i=1}^{q} \hat{\theta}_i \varepsilon_{t+l-i} \,|y_1, y_2, \cdots, y_t\right) \\ &= \begin{cases} \sum\limits_{i=1}^{l-1} \hat{\phi}_i \hat{y}_t(l-i) + \sum\limits_{i=l}^{p} \hat{\phi}_i y_{t+l-i} + \sum\limits_{i=1}^{q} \hat{\theta}_i \varepsilon_{t+l-i}, & 当 l < \min\{p,\ q\}, \\ \sum\limits_{i=1}^{p-1} \hat{\phi}_i \hat{y}_t(l-i)\,, & 当 l > \max\{p,\ q\}, \\ \sum\limits_{i=1}^{p} \hat{\phi}_i \hat{y}_t(l-i) + \sum\limits_{i=1}^{q} \hat{\theta}_i \varepsilon_{t+l-i}, & 当 p \leqslant l < q, \\ \sum\limits_{i=1}^{l-1} \hat{\phi}_i \hat{y}_t(l-i) + \sum\limits_{i=l}^{p} \hat{\phi}_i y_{t+l-i}, & 当 q \leqslant l < p. \end{cases} \end{aligned} \tag{2.3.10}$$

2.4 灰色系统预测模型

灰色系统理论是由华中科技大学邓聚龙教授[51,52] 于 1982 年提出来的. 灰色系统一词由自动控制理论中的黑箱引申而来. 黑箱 (black box) 表示人们对系统的内部结构、特征全然不知, 只能通过外部的表象对其研究. 与之相反, 人们把内部结构、特征了解得清清楚楚的系统称为白色系统. 然而在现实世界中, 我们遇到的

绝大多数社会、经济和管理系统, 对其内部结构、特征的了解介于黑色系统和白色系统之间, 所以称之为灰色系统. 灰色系统就是指部分信息已知部分信息未知的系统. 这里用颜色深浅来表示系统信息完备程度.

例如, 国民经济是由各种相互联系、相互制约的经济要素组成的, 具有一定经济特征的经济系统. 其中的经济要素主要包括原材料、设备、厂房、土地、劳动力、产品、资金等物质, 同时也包含有技术和政策等方面的经济信息. 但是我们并不十分清楚国民经济系统的内部结构和运行机制, 只知道经济系统的一些外部信息, 比如一定的投入及产出情况. 另外, 经济系统也不是单独存在的, 它受自然、社会等外界环境的影响和制约, 同外界有着物质、信息的交换. 因此, 国民经济系统是一个非常复杂的、开放的巨系统, 它是部分信息已知部分信息未知的系统, 是一个典型的灰色系统.

尽管在数理统计学中有方差分析、回归分析等因素分析方法, 但是统计方法中要求大样本及其样本具有典型的概率分布, 这限制了某些统计方法的应用. 正是从这个角度说, 灰色预测具有一定的优越性.

所谓灰色预测法, 是指由于历史数据的不全面和不充分, 或某些变量尚不清楚和不确定, 使预测处于一种半明半暗的状态. 随着事件的发展, 数据的逐步积累, 一些不确定的因素逐步明确, 其预测将逐渐由暗变明. 灰色预测法是通过建立灰色预测模型 (grey model) 来进行预测的, 该模型简称为 GM 模型. GM 模型是对原始时间数列数据进行一次累加生成后用微分方程来刻画. 它可以用阶数 M 和自变量个数 N 表示, 记为 GM(M, N). 下面介绍 GM(1, 1) 模型.

2.4.1 GM(1, 1) 预测模型的基本原理

设有原始时间数列 $x^0 = \{x^0(1),\ x^0(2),\ \cdots,\ x^0(n)\}$, 对其作一次累加生成运算, 即令

$$x^1(k) = \sum_{i=1}^{k} x^0(i), \quad k = 1, 2, \cdots, n. \tag{2.4.1}$$

从而可得新的生成数列 $x^1 = \{x^1(1),\ x^1(2),\ \cdots,\ x^1(n)\}$, 新的生成数列 x^1 一般近似地服从指数规律, 因此它满足如下灰色预测的微分方程 GM(1, 1), 其白化形式为

$$\frac{\mathrm{d}x^1}{\mathrm{d}t} + ax^1 = b, \tag{2.4.2}$$

其中 a, b 为辨识参数.

为了估计参数 a, b, 可以将式 (2.4.2) 进行离散化处理得

$$\Delta(x^1(k+1)) + a\,X^1(k+1) = b, \qquad k = 1, 2, \cdots, n-1, \tag{2.4.3}$$

其中 $\Delta(x^1(k+1))$ 为生成数列 x^1 在第 $k+1$ 时刻的累减生成, 即

$$\Delta(x^1(k+1)) = x^1(k+1) - x^1(k) = x^0(k+1). \tag{2.4.4}$$

在灰色预测中, 式 (2.4.3) 中的 $X^1(k+1)$ 为 $\dfrac{\mathrm{d}x^1}{\mathrm{d}t}$ 在第 $k+1$ 时刻的背景值, 一般取其均值生成. 即

$$X^1(k+1) = \frac{1}{2}\left[x^1(k) + x^1(k+1)\right]. \tag{2.4.5}$$

将式 (2.4.4)、(2.4.5) 代入式 (2.4.3) 中有

$$\begin{cases} x^0(2) = a\left[-\dfrac{1}{2}\left(x^1(1) + x^1(2)\right)\right] + b, \\ x^0(3) = a\left[-\dfrac{1}{2}\left(x^1(2) + x^1(3)\right)\right] + b, \\ \cdots\cdots\cdots\cdots \\ x^0(n) = a\left[-\dfrac{1}{2}\left(x^1(n-1) + x^1(n)\right)\right] + b. \end{cases} \tag{2.4.6}$$

令

$\boldsymbol{B} =$

$$\begin{bmatrix} -\dfrac{1}{2}\left(x^1(1)+x^1(2)\right) & -\dfrac{1}{2}\left(x^1(2)+x^1(3)\right) & \cdots & -\dfrac{1}{2}\left(x^1(n-1)+x^1(n)\right) \\ 1 & 1 & \cdots & 1 \end{bmatrix}^{\mathrm{T}},$$

$\boldsymbol{Y} = \left(x^0(2) \quad x^0(3) \quad \cdots \quad x^0(n)\right)^{\mathrm{T}}, \quad \boldsymbol{\alpha} = (a \quad b)^{\mathrm{T}}.$

则式 (2.4.6) 可简化为如下线性模型

$$\boldsymbol{Y} = \boldsymbol{B}\,\boldsymbol{\alpha}. \tag{2.4.7}$$

由最小二乘估计方法得

$$\boldsymbol{\alpha} = \left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}. \tag{2.4.8}$$

式 (2.4.8) 估计出来的参数代入到式 (2.4.2) 的白化形式.

令 $x^1 - \dfrac{b}{a} = p$, 则有 $\dfrac{\mathrm{d}p}{p} = -a\,\mathrm{d}t$, 由分离变量法得 $p = c\mathrm{e}^{-at}$, 其中 c 为常数. 考虑到初值 $x^1(t_0) = x^0(1)$, 所以 $c = \left(x^0(1) - \dfrac{b}{a}\right)\mathrm{e}^{-a\,t_0}$, 从而有

$$\hat{x}^1(t) = \left(x^0(1) - \frac{b}{a}\right)\mathrm{e}^{-a(t-t_0)} + \frac{b}{a}. \tag{2.4.9}$$

式 (2.4.9) 就是 GM(1, 1) 模型的时间响应函数形式, 将它离散化得

$$\hat{x}^1(k+1)=\left(x^0(1)-\frac{b}{a}\right)\mathrm{e}^{-ak}+\frac{b}{a}. \tag{2.4.10}$$

对序列 $\hat{x}^1(k+1)$ 再作累减生成可进行预测. 即

$$\begin{aligned}\hat{x}^0(k+1)&=\hat{x}^1(k+1)-\hat{x}^1(k)\\&=\left(x^{(0)}(1)-\frac{b}{a}\right)[1-\mathrm{e}^a]\mathrm{e}^{-ak},\qquad k=1,2,\cdots,n,\\&\qquad\qquad n+1,\cdots.\end{aligned} \tag{2.4.11}$$

式 (2.4.11) 便是 GM(1, 1) 模型的预测的具体计算式.

2.4.2 GM(1, 1) 预测模型的检验

灰色预测的模型的检验包括残差检验、关联度检验、后验差检验三种形式. 这里只给出常用的后验差检验. 其检验步骤如下:

第一步 计算原始时间数列 $x^0=\{x^0(1),\ x^0(2),\ \cdots,\ x^0(n)\}$ 的均值和方差

$$\bar{x}=\frac{1}{n}\sum_{k=1}^{n}x^0(k),\qquad s_1^2=\frac{1}{n}\sum_{k=1}^{n}(x^0(k)-\bar{x})^2.$$

第二步 计算残差数列 $e^0=\{e^0(1),\ e^0(2),\ \cdots,\ e^0(n)\}$ 的均值 $\bar{e}$ 和方差 s_2^2

$$\bar{e}=\frac{1}{n}\sum_{k=1}^{n}e^0(k),\qquad s_2^2=\frac{1}{n}\sum_{k=1}^{n}(e^0(k)-\bar{e})^2,$$

其中 $e^0(k)=x^0(k)-\hat{x}^0(k),\ k=1,2,\cdots,n$ 为残差数列.

第三步 计算后验差比值

$$c=s_2/s_1.$$

第四步 计算小误差频率

$$p=|A|/n,$$

其中 $A=\left\{k\,\middle|\ \left|e^0(k)-\bar{e}\right|\leqslant 0.6745s_1\ ,\ k=1,2,\cdots,n\right\}$, $|A|$ 为集合 A 的个数.

第五步 根据表 2.4.1, 按照后验差比值 c 和小误差频率 p 判别预测精度等级.

表 2.4.1 预测精度等级

等级	p	c
好	>0.95	<0.35
合格	>0.80	<0.45
勉强	>0.70	<0.50
不合格	<0.70	>0.65

如果后验差检验发现 GM(1, 1) 模型预测精度等级为不合格, 那么可以进行残差修正的 GM(1, 1) 模型.

2.4.3 灰色关联度计算式及改进

设一个母因素时间数列和 m 个子因素时间数列分别为

$$X_0=(x_0(1),x_0(2),\cdots,x_0(n)),\qquad X_i=(x_i(1),x_i(2),\cdots,x_i(n)),\qquad i=1,2,\cdots,m.$$

令 $\xi_i(t)=\dfrac{\Delta_{\min}+\rho\Delta_{\max}}{\Delta+\rho\Delta_{\max}}$, 则称 $\xi_i(t)$ 表示第 t 时刻子因素和母因素的关联系数, 其中 ρ 称为分辨系数, $\rho\in[0,\ 1]$, 一般取 $\rho=0.5$.

$$\begin{aligned}\Delta&=|x_0(t)-x_i(t)|,\\ \Delta_{\min}&=\min\min|x_0(t)-x_i(t)|,\\ \Delta_{\max}&=\max\max|x_0(t)-x_i(t)|.\end{aligned}$$

则母因素和子因素的关联度 R_i 计算式[52] 定义为

$$R_i=\frac{1}{n}\sum_{t=1}^{n}\xi_i(t). \tag{2.4.12}$$

式 (2.4.12) 的关联度计算式存在如下问题:

(1) R_i 与分辨系数有关, R_i 的值不唯一.

(2) 灰关联度不具有保序性. 即 R_i 不是 ρ 的单调函数. 对某个 ρ_0, 可能 $R_i>R_j$, 而对另外一个 ρ_0', 可能 $R_i<R_j$.

(3) 关联度不具有规范性. 即对完全相关的子因素与母因素的时间数列, $R_i\neq 1$. 例如, $X_0=(1,2,\cdots,10)$, $X_1=(11,12,\cdots,20)$, 显然 X_0 与 X_1 完全相关, 但是若取 $\rho=0.5$, 则按上式算 $R_1=1/3$.

(4) 灰关联分析不能体现负相关. 由灰关联度计算式知 $R_i>0$. 然而, 设有两个时间数列 $X_0=(1,2,\cdots,20)$, $X_1=(20,\ 19,\cdots,1)$, 显然有 $x_1(t)=21-x_0(t)$, $t=1,2,\cdots,20$, X_0 与 X_1 存在典型的负相关.

(5) 通常分辨系数 $\rho=0.5$, 则有灰关联度的值恒大于或等于 $\dfrac{1}{3}$.

鉴于上述关联度计算式存在的问题, 有必要提出改进的关联度计算式. 文献 [53] 提出改进的关联度计算式, 步骤如下:

(1) 作一次累减生成, 即相当于求所在曲线上不同时点的斜率.

$$a^{(1)}(x_0(t))=x_0(t+1)-x_0(t),\qquad a^{(1)}(x_i(t))=x_i(t+1)-x_i(t),\qquad t=1,2,\cdots,n-1.$$

(2) 计算 X_0 与 X_i 两时间数列的标准差. 记

$$\bar{x}_0 = \frac{1}{n}\sum_{t=1}^{n} x_0(t) \qquad \bar{x}_i = \frac{1}{n}\sum_{t=1}^{n} x_i(t)\ ,$$

$$\sigma_{x_0} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(x_0(t)-\bar{x}_0)^2}, \qquad \sigma_{x_i} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(x_i(t)-\bar{x}_i)^2}.$$

即 $\bar{x}_0$ 和 $\bar{x}_i$ 分别表示两时间数列 X_0 和 X_i 的均值, $\sigma_{x_0}, \sigma_{x_i}$ 分别表示两时间数列 X_0 和 X_i 的标准差.

(3) 计算 t 时刻 $x_0(t)$, $x_i(t)$ 的关联系数.

$$\xi(x_0(t), x_i(t)) = \mathrm{sign}(a^{(1)}(x_0(t))a^{(1)}(x_i(t)))\frac{1}{1+\left|\left|\dfrac{a^{(1)}(x_0(t))}{\sigma_{x_0}}\right| - \left|\dfrac{a^{(1)}(x_i(t))}{\sigma_{x_i}}\right|\right|},$$

其中

$$\mathrm{sign}x = \begin{cases} 1, & 当x > 0, \\ 0, & 当x = 0, \\ -1, & 当x < 0. \end{cases}$$

(4) 计算 X_0 与 X_i 的灰关联度

$$\tilde{R}_i = \frac{1}{n-1}\sum_{t=1}^{n-1}\xi(x_0(t), x_i(t)). \tag{2.4.13}$$

上述改进的关联度式 (2.4.13) 具有唯一性和规范性 [53], 即

$$|\tilde{R}_i| \leqslant 1且\tilde{R}_i = 1当且仅当X_0与X_i完全相关.$$

2.5 季节变动时间序列的预测模型

由于受到自然条件和社会条件的影响, 在一个年度内某些社会经济现象出现季节性的变化. 所有这些因季节性改变而带有规律性的变动称为季节变动. 如某些农产品的产量、重要节日期间交通运输部门的客运量等. 对季节变动的时间序列进行预测方法的研究可以克服季节变动所带来的不良影响, 有利于人们安排生产和生活, 具有重要的现实意义. 季节变动的时间序列的预测模型大体上可分为两类: 一类是加法型的预测模型, 另一类是乘法型的预测模型. 这里介绍乘法型的预测模型.

2.5.1 季节变动时间序列乘法型预测模型

某社会经济现象的季节变动时间序列为 $\{x_t, t=1,2,\cdots,N\}$，设 T_t 表示趋势变化值，$S_t>0$ 表示季节指数，$\sum\limits_{t=1}^{k} S_t = k$，$k=4$ 表示时间序列 $\{x_t, t=1,2,\cdots,N\}$ 为季度数据，$k=12$ 表示时间序列 $\{x_t, t=1,2,\cdots,N\}$ 为月度数据，ε_t 表示相互独立的随机季节变动因子，$\varepsilon_t \sim N(0,\sigma^2)$，乘法型的预测模型可表示为

$$x_t = T_t S_t \mathrm{e}^{\varepsilon_t}, \tag{2.5.1}$$

利用乘法型模型进行预测的步骤为：

第一步　时间序列 $\{x_t, t=1,2,\cdots,N\}$ 分解为长期趋势 $\hat{T}_t = \hat{a} + \hat{b}t$，其中 $\hat{T}_t$ 是根据滑动平均值序列

$$\mathrm{MA}(x_t) = \frac{x_t + x_{t-1} + \cdots + x_{t-k+1}}{k},$$

其中 k 表示时间序列滑动平均项数.

第二步　利用序列 $\{x_t/\mathrm{MA}(x_t)\}$ 分解求出季节指数. 由于 ε_t 表示相互独立的随机季节变动因子，且 $E\varepsilon_t = 0$，所以可采用平均的方法消除随机性影响. 将序列 $\{x_t/\mathrm{MA}(x_t)\}$ 按季节排列，再把各年相同季节相加起来进行平均即可分解季节指数 $\hat{S}_l, l=1,2,\cdots,k$.

第三步　根据前两步计算的长期趋势 $\hat{T}_t = \hat{a} + \hat{b}t$ 和季节指数 $\hat{S}_l, l=1,2,\cdots,k$，按下式进行预测

$$\hat{x}_{n+T} = \hat{T}_{n+T}\hat{S}_l, \tag{2.5.2}$$

其中 $n+T = l+km$，m 为整数.

2.5.2 季节变动时间序列乘法型渐近预测模型

设共有 n 年 m 个季节的某社会经济现象的季节变动时间序列为 $\{x_t, t=1,2,\cdots,N\}$，其中 $N=nm$，将其按时间进行如下下标转换

$$x_{ij} = x_{(i-1)m+j}, \qquad i=1,2,\cdots,n, j=1,2,\cdots,m,$$

其中 x_{ij} 表示第 i 年 j 个季节的实际观察值，m 表示一年包含的季节数. 即若 $\{x_t, t=1,2,\cdots,N\}$ 是月度资料，则 m=12；若 $\{x_t, t=1,2,\cdots,N\}$ 是季度资料，则 m=4. 称矩阵 $\boldsymbol{X} = (x_{ij})_{n\times m}$ 为季节变动时间序列 $\{x_t, t=1,2,\cdots,N\}$ 所对应的矩阵.

令

$$\bar{X} = \sqrt[nm]{\prod_{i=1}^{n}\prod_{j=1}^{m} x_{ij}}, \qquad \bar{X}_{i\cdot} = \sqrt[m]{\prod_{j=1}^{m} x_{ij}}, \quad \bar{X}_{\cdot j} = \sqrt[n]{\prod_{i=1}^{n} x_{ij}}, \tag{2.5.3}$$

其中 $i=1,2,\cdots,n,\ j=1,2,\cdots,m$. 则称 $\bar{X}$ 为季节变动时间序列的总平均值, 称 $\bar{X}_{i\cdot}$ 为第 i 年的年度平均值, 称 $\bar{X}_{\cdot j}$ 为第 j 个季节的季度平均值.

令

$$k_i=\bar{X}_{i\cdot}/\bar{X},\qquad l_j=\bar{X}_{\cdot j}/\bar{X},$$

$$\varepsilon_{ij}=x_{ij}/(\bar{X}k_il_j),\qquad i=1,2,\cdots,n,\qquad j=1,2,\cdots,m. \tag{2.5.4}$$

则称 k_i 是季节变动时间序列第 i 年的年度变动因子, l_j 是第 j 个季节的季节变动因子, 称 ε_{ij} 是第 i 年第 j 个季节的随机变动因子.

由式 (2.5.4) 中 ε_{ij} 的定义知

$$x_{ij}=k_il_j\bar{X}\,\varepsilon_{ij},\qquad i=1,2,\cdots,n,\qquad j=1,2,\cdots,m. \tag{2.5.5}$$

式 (2.5.5) 表明, 季节变动时间序列可以分解成年度变动因子、季节变动因子、总平均值和随机变动因子的乘积, 它就是本节提出的季节变动时间序列的一种乘法分解模型. 此乘法分解模型不同于传统的时间序列的乘法分解式 (2.5.1).

由式 (2.5.3)~(2.5.5) 易得

性质 2.5.1 季节变动时间序列的年度变动因子、季节变动因子和随机变动因子满足如下关系式

$$\prod_{i=1}^{n}k_i=1,\qquad \prod_{j=1}^{m}l_j=1,\qquad \prod_{i=1}^{n}\prod_{j=1}^{m}\varepsilon_{ij}=1. \tag{2.5.6}$$

性质 2.5.1 表明, 季节变动时间序列的 n 个年度变动因子之积为 1, m 个季节变动因子之积为 1, mn 个随机变动因子之积为 1.

在季节变动时间序列的预测中, 一般均假设随机变动因子对季节变动时间序列的预测值的影响是微不足道的, 即在本模型中, 第 i 年第 j 个季节的随机变动因子恒为 1.

对于季节变动时间序列所对应的矩阵 $\boldsymbol{X}=(x_{ij})_{n\times m}$, 假设只有 x_{nm} 是未知的, 在随机变动因子恒为 1 的情况下, 如下的性质 2.5.2 给出了利用矩阵中已知的 $(nm-1)$ 个季节变动时间序列数据来预测 x_{nm} 的渐近预测方法.

性质 2.5.2 令 $A=\prod_{j=1}^{m-1}x_{nj}$, $B=\prod_{i=1}^{n-1}x_{im}$, $C=\prod_{i=1}^{n}\prod_{j=1}^{m}x_{ij}/x_{nm}$, 设随机变动因子 $\varepsilon_{ij}=1, i=1,2,\cdots,n,\ j=1,2,\cdots,m$, 则有 x_{nm} 的预测值为

$$x_{nm}=\sqrt[(n-1)(m-1)]{A^nB^m/C}. \tag{2.5.7}$$

证明 因为 $\varepsilon_{ij}=1$, 则式 (2.5.4) 就化成

$$x_{ij}=k_il_j\bar{X},\quad i=1,2,\cdots,n, j=1,2,\cdots,m. \tag{2.5.8}$$

将上式代入到表达式 A,B,C 中, 并注意到 $\prod_{j=1}^{m} l_j = 1$, $\prod_{i=1}^{n} k_i = 1$, 有

$$A = \prod_{j=1}^{m-1} x_{nj} = \prod_{j=1}^{m-1} (k_n l_j \bar{X}) = (k_n \bar{X})^{(m-1)} \left(\prod_{j=1}^{m} l_j / l_m\right) = (k_n \bar{X})^{(m-1)} / l_m,$$

同理有

$$B = \prod_{i=1}^{n-1} x_{im} = \prod_{i=1}^{n-1} (k_i l_m \bar{X}) = (l_m \bar{X})^{(n-1)} \left(\prod_{i=1}^{n} k_i / k_n\right) = (l_m \bar{X})^{(n-1)} / k_n,$$

$$\begin{aligned} C &= \prod_{i=1}^{n} \prod_{j=1}^{m} x_{ij} / x_{nm} = \prod_{i=1}^{n} \prod_{j=1}^{m} (k_i l_j \bar{X}) / (k_n l_m \bar{X}) \\ &= \bar{X}^{nm} \left(\prod_{i=1}^{n} k_i \prod_{j=1}^{m} l_j / (k_n l_m \bar{X})\right) = \bar{X}^{(nm-1)} / (k_n l_m). \end{aligned}$$

所以

$$\begin{aligned} &\sqrt[(n-1)(m-1)]{A^n B^m / C} \\ &= \sqrt[(n-1)(m-1)]{\left[(k_n \bar{X})^{(m-1)} / l_m\right]^n \left[(l_m \bar{X})^{(n-1)} / k_n\right]^m / \left[\bar{X}^{(nm-1)} / (k_n l_m)\right]} \\ &= k_n l_m \bar{X}. \end{aligned}$$

又由式 (2.5.8) 知 $x_{nm} = k_n l_m \bar{X}$, 所以 x_{nm} 的估计值为

$$x_{nm} = \sqrt[(n-1)(m-1)]{A^n B^m / C}.$$

性质 2.5.2 表明, 在季节变动的时间序列的随机变动因子恒为 1 的条件下, 若将季节变动的时间序列进行适当的循环变动, 即不断去掉一个离当前时间最远的历史数据, 并补充一个新的预测值数据来进行滚动预测. 式 (2.5.7) 就是渐近预测方法.

另外, 根据式 (2.5.6), 还可以给出季节变动时间序列的一个乘法型的预测模型, 为此要对预测年份的年度变动因子进行估计.

性质 2.5.3　设随机变动因子 $\varepsilon_{ij} = 1, i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, m$, 季节变动时间序列的第 $(n+1)$ 个年度变动因子 k_{n+1} 的估计为

$$k_{n+1} = \sqrt[(n-1)]{k_n^n / k_1}.$$

证明　将季节变动时间序列 $\{x_t, t = 1, 2, \cdots, N\}$ 所对应的矩阵 $\boldsymbol{X} = (x_{ij})_{n \times m}$ 作一步循环变动, 即去掉 x_{11}, 补充 $x_{(n+1)\,1}$, 其他顺序不变, 其中 $x_{(n+1)\,1}$ 是未知的, 是要预测的变量. 因为 $\varepsilon_{ij} = 1$, 则式 (2.5.8) 成立, 即

$$x_{ij} = k_i l_j \bar{X}, \qquad i = 1, 2, \cdots, n, \qquad j = 1, 2, \cdots, m.$$

从而一步循环变动矩阵就变成

$$\begin{pmatrix} x_{12} & x_{13} & \cdots & x_{1m} & x_{21} \\ x_{22} & x_{23} & \cdots & x_{2m} & x_{31} \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n2} & x_{n3} & \cdots & x_{nm} & x_{(n+1)1} \end{pmatrix} = \begin{pmatrix} k_1l_2\bar{X} & k_1l_3\bar{X} & \cdots & k_1l_m\bar{X} & k_2l_1\bar{X} \\ k_2l_2\bar{X} & k_2l_3\bar{X} & \cdots & k_2l_m\bar{X}\ C & k_3l_1\bar{X} \\ \vdots & \vdots & & \vdots & \vdots \\ k_nl_2\bar{X} & k_nl_3\bar{X} & \cdots & k_nl_m\bar{X} & x_{(n+1)1} \end{pmatrix}. \tag{2.5.9}$$

由式 (2.5.8), 按性质 2.5.2 中表达式 A,B,C 的定义计算得

$$A = \left(k_n\bar{X}\right)^{(m-1)}\prod_{j=2}^{m} l_j, \quad B = \left(l_1\bar{X}\right)^{(n-1)}\prod_{i=2}^{n} k_i,$$
$$C = \left(\bar{X}\right)^{(mn-1)}\prod_{i=1}^{n} k_i^m \prod_{j=1}^{m} l_j^n / l_1k_1.$$

由式 (2.5.6) 知

$$x_{(n+1)\ 1} = \sqrt[(n-1)(m-1)]{A^nB^m/C} = \sqrt[(n-1)]{k_n^n/k_1}\ l_1\bar{X}\ . \tag{2.5.10}$$

又由式 (2.5.7) 知

$$x_{(n+1)\ 1} = k_{n+1}l_1\bar{X}. \tag{2.5.11}$$

比较 (2.5.10)、(2.5.11) 两式得 $k_{n+1} = \sqrt[(n-1)]{k_n^n/k_1}$.

性质 2.5.3 给出了季节变动时间序列的第 $(n+1)$ 个年度变动因子估计式. 在季节变动的时间序列的随机变动因子恒为 1 的条件下, 由季节变动时间序列新的乘法分解模型式 (2.5.5) 可以得出第 $n+1$ 个年度第 j 个季节的预测值的另一个估计式为

$$x_{(n+1)\ j} = k_{n+1}l_j\bar{X} = \sqrt[(n-1)]{k_n^n/k_1}\ l_j\bar{X}, \quad j = 1, 2, \cdots, m, \tag{2.5.12}$$

式 (2.5.12) 就是乘法型的预测方法.

2.5.3 实例分析

应用文献 [1] 的数据进行实例分析. 某地 1991~1996 年某种服装销售量有下述销售数据 (见文献 [1] 第 133 页), 如表 2.5.1 所示.

文献 [1] 绘制出销售量依时间变化的图形, 从图形可以明显地看出该服装销售量的时间序列具有季节变动性, 下面利用本书给出的两种预测法来预测 1997 年各个季节的服装销售量.

表 2.5.1　某地某种服装销售量表 (单位：万件)

年份t	第一季度销售量	第二季度销售量	第三季度销售量	第四季度销售量
1991	80	70	90	100
1992	90	80	105	120
1993	98	90	110	130
1994	104	100	120	140
1995	114	104	130	148
1996	122	112	138	158

方法一　渐近预测法.

要预测 1997 年第一季度服装销售量 $\hat{x}_{71}$, 先去掉 x_{11}, 补充 $\hat{x}_{71}$, 其他顺序不变, 得一步循环变动矩阵, 其中 $\hat{x}_{71}$ 是未知的.

$$\begin{pmatrix} x_{12} & x_{13} & x_{14} & x_{21} \\ x_{22} & x_{23} & x_{24} & x_{31} \\ x_{32} & x_{33} & x_{34} & x_{41} \\ x_{42} & x_{43} & x_{44} & x_{51} \\ x_{52} & x_{53} & x_{54} & x_{61} \\ x_{62} & x_{63} & x_{64} & \hat{x}_{71} \end{pmatrix} = \begin{pmatrix} 70 & 90 & 100 & 90 \\ 80 & 105 & 120 & 98 \\ 90 & 110 & 130 & 104 \\ 100 & 120 & 140 & 114 \\ 104 & 130 & 148 & 122 \\ 112 & 138 & 158 & \hat{x}_{71} \end{pmatrix}. \tag{2.5.13}$$

利用 Matlab 软件计算得

$$A = 112 \times 138 \times 158 = 2.4420 \times 10^6,$$

$$B = 90 \times 98 \times 104 \times 114 \times 122 = 1.2758 \times 10^{10},$$

$$C = 8.5595 \times 10^{46}(\text{式 (2.5.13) 中除 } \hat{x}_{71}\text{之外的其他元素的乘积}).$$

将其代入渐近预测公式 (2.5.7) 得

$$\hat{x}_{71} = \sqrt[15]{A^6 B^4 / C} = 132.17(\text{万件}).$$

为了预测 1997 年第二季度服装销售量 $\hat{x}_{72}$, 将季节变动时间序列去掉 x_{12}, 补充 $\hat{x}_{71}$, 其他顺序不变, 得一步循环变动矩阵

$$\begin{pmatrix} x_{13} & x_{14} & x_{21} & x_{22} \\ x_{23} & x_{24} & x_{31} & x_{32} \\ x_{33} & x_{34} & x_{41} & x_{42} \\ x_{43} & x_{44} & x_{51} & x_{52} \\ x_{53} & x_{54} & x_{61} & x_{62} \\ x_{63} & x_{64} & \hat{x}_{71} & \hat{x}_{72} \end{pmatrix} = \begin{pmatrix} 90 & 100 & 90 & 80 \\ 105 & 120 & 98 & 90 \\ 110 & 130 & 104 & 100 \\ 120 & 140 & 114 & 104 \\ 130 & 148 & 122 & 112 \\ 138 & 158 & 132.17 & \hat{x}_{72} \end{pmatrix}.$$

按渐近预测方法公式 (2.5.7) 计算得

$$\hat{x}_{72} = 121.04(\text{万件}), \qquad \hat{x}_{73} = 150.46(\text{万件}), \qquad \hat{x}_{74} = 173.97(\text{万件}).$$

方法二 乘法预测模型.

根据表 2.5.1 销售数据计算得

$$\bar{X} = \sqrt[24]{\prod_{i=1}^{6}\prod_{j=1}^{4} x_{ij}} = 108.3463, \quad k_1 = \bar{X}_1./\bar{X} = 0.7777, \quad k_2 = \bar{X}_2./\bar{X} = 0.9008,$$

$$k_3 = \bar{X}_3./\bar{X} = 0.9781, \qquad k_4 = \bar{X}_4./\bar{X} = 1.0611, \qquad k_5 = \bar{X}_5./\bar{X} = 1.1343,$$

$$k_6 = \bar{X}_6./\bar{X} = 1.2126, \qquad l_1 = \bar{X}._1/\bar{X} = 0.9261, \qquad l_2 = \bar{X}._2/\bar{X} = 0.8446,$$

$$l_3 = \bar{X}._3/\bar{X} = 1.0556, \qquad l_4 = \bar{X}._4/\bar{X} = 1.2112.$$

按乘法型的预测模型公式 (2.5.12) 计算得

$$\hat{x}_{71} = \sqrt[5]{k_6^6/k_1}\, l_1\bar{X} = 132.97(\text{万件}), \qquad \hat{x}_{72} = \sqrt[5]{k_6^6/k_1}\, l_2\bar{X} = 121.27(\text{万件}),$$

$$\hat{x}_{73} = \sqrt[5]{k_6^6/k_1}\, l_3\bar{X} = 151.56(\text{万件}), \qquad \hat{x}_{74} = \sqrt[5]{k_6^6/k_1}\, l_4\bar{X} = 173.91(\text{万件}).$$

文献 [1] 利用传统的时间序列的乘法分解预测 1997 年各个季节的服装销售量的结果如下：

$$\hat{x}_{71} = 129.29(\text{万件}), \qquad \hat{x}_{72} = 119.28(\text{万件}),$$

$$\hat{x}_{73} = 147.36(\text{万件}), \qquad \hat{x}_{74} = 168.66(\text{万件}).$$

从上面预测结果来看, 上述提出两种预测方法所得结果相差无几, 且和文献 [1] 所得结果基本一致, 从而表明上述提出两种预测方法的有效性.

第 3 章 非最优的组合预测模型

3.1 组合预测的分类

组合预测集结各单项预测方法的特点, 可以从不同的角度进行分类. 根据其目标和特点不同, 大体上可从如下几个角度分类:

(1) 按组合预测与各单项预测方法的函数关系, 组合预测可以分成线性组合预测和非线性组合预测.

设预测对象存在 m 个单项预测方法, 利用这 m 个单项预测方法得到的第 i 个单项预测方法的预测值为 f_i, $i=1,2,\cdots,m$.

若组合预测值 f 满足 $f=l_1f_1+l_2f_2+\cdots+l_mf_m$, 则称该组合预测为线性组合预测, 其中 $l_1,l_2,\cdots,l_m$ 为各种预测方法的加权系数, 一般 $\sum\limits_{i=1}^{m}l_i=1$, $l_i\geqslant 0$, $i=1,2,\cdots,m$.

若组合预测值 f 满足 $f=\phi(f_1,f_2,\cdots,f_m)$, 其中 ϕ 为非线性函数, 则称该组合预测为非线性组合预测.

常见的非线性组合预测形式有

$$f=\prod_{i=1}^{m}f_i^{l_i}.$$

上式称为加权几何平均组合预测模型.

$$f=\frac{1}{\sum\limits_{i=1}^{m}\dfrac{l_i}{f_i}}.$$

此式称为加权调和平均组合预测模型.

(2) 按组合预测加权系数计算方法的不同, 组合预测方法可以分为最优组合预测方法和非最优组合预测方法.

最优组合预测方法的基本思想就是根据某种准则构造目标函数, 在一定的约束条件下求得目标函数的最大值或最小值, 从而求得组合预测方法加权系数. 最优组

合预测方法一般可以表示成如下数学规划问题

$$\max(\min)\ Q = Q(l_1, l_2, \cdots, l_m),$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i = 1\ , \\ l_i \geqslant 0,\ i = 1, 2, \cdots, m, \end{cases}$$

其中 $Q(l_1, l_2, \cdots, l_m)$ 为目标函数, $l_1, l_2, \cdots, l_m$ 为各种单项预测方法加权系数, 加权系数可以考虑允许负和非负两种情形.

在求解一些最优组合预测模型时可能出现组合预测的权系数为负的现象[54], 而负的组合预测的权系数没有实际的意义. 非最优正权组合预测方法正好可以克服这个不足之处.

非最优正权组合预测方法就是根据预测学的基本原理, 并力求简便的原则来确定组合预测的权系数的一种方法. 具体地说, 就是根据各个单项预测模型预测的误差的方差和其权系数成反比的基本原理, 给出组合预测的权系数的计算公式. 显然, 非最优正权组合预测方法目标函数值一般要劣于最优正权组合预测方法目标函数值.

(3) 按组合预测加权系数是否随时间变化, 组合预测方法可以分为不变权组合预测方法和可变权组合预测方法.

不变权组合预测方法就是通过最优化规划模型或其他方法计算出各个单项预测方法在组合预测中的权系数. 假定它们不变, 并用这个权系数进行预测. 然而在预测实践中, 就每一个单项预测方法而言, 它经常出现对同一预测对象的不同时间上预测精度的不一致性, 也就是说, 有些时点上预测精度好, 有些时点上预测精度差. 所以不变权组合预测方法显然没有可变权组合预测方法科学.

所谓可变权组合预测方法就是组合预测加权系数随时间变化而变化. 目前可变权组合预测方法比较复杂, 因此可变权组合预测方法的研究成果并不多见. 变权组合预测方法有待于进一步研究, 这也是组合预测方法今后重要的研究方向之一.

(4) 从某个准则的结果优劣程度来看, 组合预测方法可以分为非劣性组合预测和优性组合预测 [45].

按某个准则, 把组合预测的结果和各个单项预测方法的结果进行对比. 若组合预测的结果介于各个单项预测方法结果“最差”和“最好”之间, 则称该组合预测为非劣性组合预测. 若组合预测的结果比各个单项预测方法结果“最好”的还要“好”, 则称该组合预测为优性组合预测.

显然, 组合预测方法是建立在充分利用已知信息基础上的, 它集结各个单项预测方法所包含的信息进行组合. 所以只有当组合预测为优性组合预测时, 组合预测方法才有实际的意义. 也就是说, 通过组合预测可以达到提高预测精度、改善预测

结果的目的.

3.2 非最优正权组合预测模型权系数的确定方法

3.2.1 几种常规的非最优正权组合预测模型权系数的确定方法

组合预测的核心问题就是如何求出加权平均系数, 使得组合预测模型更加有效地提高预测精度. 若以预测绝对误差作为预测精度的衡量指标, 则主要有几种常规的非最优正权组合预测模型权系数的确定方法.

1. *算术平均方法*

即令

$$l_i = \frac{1}{m}, \qquad i = 1, 2, \cdots, m. \tag{3.2.1}$$

显然 $\sum_{i=1}^{m} l_i = 1$, $l_i \geqslant 0$, $i = 1, 2, \cdots, m$.

算术平均方法也称为等权平均方法. 算术平均方法的特点是 m 种单项预测方法的加权系数完全相等, 即把各个单项预测模型同等看待. 算术平均方法一般使用在对各个单项预测模型的预测精度缺乏了解的情形. 由于算术平均方法的计算简单, 且加权系数也满足非负性, 所以它在预测领域中的应用比较广泛. 当各个单项预测模型的预测精度完全已知时, 一般要采用加权平均的形式, 对预测精度高的单项预测模型的应赋予较大的加权系数. 下面的几种方法体现这个特点.

2. *预测误差平方和倒数方法*

预测误差平方和倒数方法也称为方差倒数方法, 这是对等权平均方法的改进. 一般说来每种单项预测模型的预测精度不同. 预测误差平方和是反映预测精度的一个指标. 预测误差平方和越大, 表明该项预测模型的预测精度就越低, 从而它在组合预测中的重要性就降低. 重要性的降低表现为它在组合预测中的加权系数就越小. 反之, 对预测误差平方和较小的单项预测模型在组合预测中的应赋予较大的加权系数. 令

$$l_i = \frac{E_{ii}^{-1}}{\sum_{i=1}^{m} E_{ii}^{-1}}, \qquad i = 1, 2, \cdots, m. \tag{3.2.2}$$

显然 $\sum_{i=1}^{m} l_i = 1$, $l_i \geqslant 0$, $i = 1, 2, \cdots, m$, 其中 E_{ii} 为第 i 种单项预测模型的预测误差

平方和

$$E_{ii}=\sum_{t=1}^{N}e_{it}^2=\sum_{t=1}^{N}(x_t-x_{it})^2. \tag{3.2.3}$$

x_{it} 为第 i 种单项预测方法在第 t 时刻的预测值, x_t 为同一预测对象的某个指标序列 $\{x_t\ ,t=1,2,\cdots,N\}$ 第 t 时刻的观测值. N 表示时间长度. $e_{it}=(x_t-x_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差.

3. 均方误差倒数方法

均方误差倒数方法的含义类似于预测误差平方和倒数方法. 该方法体现了某单项预测模型的误差平方和越大, 它在组合预测中的加权系数就应越小. 均方误差倒数方法的加权系数的计算公式为

$$l_i=\frac{E_{ii}^{-\frac{1}{2}}}{\sum\limits_{i=1}^{m}E_{ii}^{-\frac{1}{2}}},\qquad i=1,2,\cdots,m. \tag{3.2.4}$$

显然 $\sum\limits_{i=1}^{m}l_i=1$, $l_i\geqslant 0$, $i=1,2,\cdots,m$, 其中 E_{ii} 的含义同上.

4. 简单加权平均方法

简单加权平均方法也是一种非等权平均方法. 它是先把各个单项预测模型预测的误差的方差和 E_{ii}, $i=1,2,\cdots,m$ 进行排序, 不妨设 $E_{11}>E_{22}>\cdots>E_{mm}$, 根据各个单项预测模型预测的误差的方差和的权系数成反比的基本原理知, 排序越靠前面的单项预测模型, 在组合预测中的加权系数就应越小. 即令

$$l_i=\frac{i}{\sum\limits_{i=1}^{m}i}=\frac{2i}{m(m+1)},\qquad i=1,2,\cdots,m. \tag{3.2.5}$$

显然 $\sum\limits_{i=1}^{m}l_i=1$, $l_i\geqslant 0$, $i=1,2,\cdots,m$, 其中 E_{ii} 的含义同上.

5. 二项式系数方法

二项式系数方法和简单加权平均方法有一点相似之处, 它也是先把各个单项预测模型预测的误差的方差和 E_{ii}, $i=1,2,\cdots,m$ 进行排序, 不妨设 $E_{11}>E_{22}>\cdots>E_{mm}$, 但它取组合预测中的加权系数的思想和简单加权平均方法是不同的. 它是按着统计学的中位数的概念, 若单项预测模型预测的误差的方差和过大或过小, 则其对应的权系数均较小, 而处于各单项预测模型预测的误差的方差和的中位数所对应

的权系数最大. 即令

$$l_i = \frac{C_{2m-1}^{i-1}}{2^{2m-2}}, \quad i = 0, 1, 2, \cdots, m-1. \tag{3.2.6}$$

由二项式定理知

$$1 = \left(\frac{1}{2} + \frac{1}{2}\right)^{2m-1} = \sum_{i=0}^{2m-1} C_{2m-1}^{i} \left(\frac{1}{2}\right)^{2m-1}.$$

且 $C_{2m-1}^{i} = C_{2m-1}^{2m-1-i}$, 则有 $\sum\limits_{i=0}^{m-1} C_{2m-1}^{i} \left(\frac{1}{2}\right)^{2m-1} = \frac{1}{2}$, 即

$$\sum_{i=0}^{m-1} C_{2m-1}^{i} \left(\frac{1}{2}\right)^{2m-2} = 1.$$

所以 $\sum\limits_{i=1}^{m} l_i = 1$, $l_i \geqslant 0$, $i = 1, 2, \cdots, m$.

3.2.2 非最优组合预测系数确定方法的应用举例

太阳黑子数多种时间序列模型的组合预测. 对太阳活动规律的模拟与预测是全球多学科研究的热点, 尤其是太阳黑子相对数极大极小的年份和数值更为世界所关注. 已用太阳黑子数 1700~1979 年的数据建立了五种时间序列模型. 表 3.2.1 是 1980~1987 年太阳黑子数的观察值和五种时间序列模型的预测值[55].

表 3.2.1 太阳黑子数和五种时间序列模型的预测值

年份	黑子数的观察值	TAR 模型	含趋势叠合模型	不含趋势叠合模型	ARMA 模型	AR 模型
1980	154.6	158.63	167.5	163.5	167.7	158.97
1981	140.4	137.24	147.5	139.5	140.8	128.99
1982	115.9	98.21	110.7	100.4	94.47	84.74
1983	66.6	61.31	72.21	65.38	48.42	48.64
1984	45.9	32.64	48.88	35.02	16.47	21.66
1985	17.9	17.42	20.77	15.2	4.05	9.38
1986	13.4	16.4	12.11	7.812	9.09	10.7
1987	29.2	26.47	26.3	21.75	24.88	32.23

表 3.2.1 的结果显示, 五种时间序列模型的预测值在 1980~1987 年太阳黑子数有基本一致的起伏规律. 即均在 1980 年太阳黑子数达到高峰, 在 1987 年太阳黑子数跌到谷底. 在具体数值上五种时间序列模型有较大的差异. 下面用上述的五种权系数的确定方法给出其组合预测模型, 并作对比分析.

(1) 算术平均方法, 得出的组合预测权系数向量为

$$(l_1, l_2, l_3, l_4, l_5) = (0.2, 0.2, 0.2, 0.2, 0.2).$$

(2) 预测误差平方和倒数方法, 得出的组合预测权系数向量为

$$(l_1, l_2, l_3, l_4, l_5) = (0.2255, 0.4172, 0.2363, 0.0614, 0.0596).$$

(3) 均方误差倒数方法, 得出的组合预测权系数向量为

$$(l_1, l_2, l_3, l_4, l_5) = (0.2263, 0.3078, 0.2316, 0.1180, 0.1163).$$

(4) 简单加权平均方法, 得出的组合预测权系数向量为

$$(l_1, l_2, l_3, l_4, l_5) = (3/15, 5/15, 4/15, 2/15, 1/15).$$

(5) 二项式系数方法, 得出的组合预测权系数向量为

$$(l_1, l_2, l_3, l_4, l_5) = (0.1406, 0.4922, 0.3281, 0.0352, 0.0039).$$

五种组合预测模型对 1980～1987 年太阳黑子数的组合预测值和预测精度见表 3.2.2.

表 3.2.2 1980～1987 年太阳黑子数的组合预测值和预测精度

年 份	方法 (1)	方法 (2)	方法 (3)	方法 (4)	方法 (5)
1980	163.26	164.06	163.60	164.12	164.91
1981	138.81	141.78	140.38	141.19	143.12
1982	97.70	102.91	100.55	101.56	104.89
1983	59.19	65.27	62.61	63.46	67.51
1984	30.93	38.33	35.00	35.80	40.80
1985	13.36	16.99	15.42	15.63	17.84
1986	11.22	11.79	11.56	11.33	11.19
1987	26.33	25.53	25.81	25.33	24.80
平均绝对误差	7.55	4.86	5.87	5.76	4.59
误差平方和	721.07	336.10	472.11	433.13	285.99

从表 3.2.2 可以看出, 以平均绝对误差和误差平方和作为组合预测精度的两个基本指标, 在上述的五种权系数的确定方法中, 二项式系数方法是最好的组合预测方法, 其次是预测误差平方和倒数方法、简单加权平均方法、均方误差倒数方法, 算术平均方法是较差的组合预测方法.

3.3 组合预测权系数确定的一种合作对策方法

组合预测权系数确定的方法已有多种, 本节在这方面作进一步的探索. 即从对策论的观点出发, 视各单项预测方法为组合预测方法这个合作对策的局中人, 合作的"结果"为组合预测的误差平方和, 再按合作对策 Shapley 值法在各单项预测模型中进行分配, 从而获得组合预测权系数确定的一种方法. 实例计算结果令人满意.

3.3.1 组合预测方法的合作对策描述 [56]

设对同一预测对象的某个指标序列 $\{x_t, t=1,2,\cdots,N\}$, 存在 m 种单项预测方法对其进行预测, m 种单项预测方法表示为 $M=\{1,2,\cdots,m\}$, 则 M 为组合预测方法的局中人集合.

记 M 的所有的子集为 2^M, 则 M 中的任一子集 $s\in 2^M$ 形成组合预测方法的一个联盟, 若干个局中人结成联盟后, 这个联盟作为一个整体进行组合预测就是希望尽可能多地降低组合预测误差. 下面采用组合预测误差平方和这个指标来反映预测精度.

设第 i 种单项预测方法在第 t 时刻的预测值为 x_{it}, $i=1,2,\cdots,m$, $t=1,2,\cdots,N$, 称 $e_{it}=(x_t-x_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差, 设 $l_1,l_2,\cdots,l_m$ 分别为 m 种单项预测方法的加权系数, $\hat{x}_t=l_1x_{1t}+l_2x_{2t}+\cdots+l_mx_{mt}$为$x_t$ 的组合预测值, 加权系数应满足

$$l_1+l_2+\cdots+l_m=1,\qquad l_1,l_2,\cdots,l_m\geqslant 0. \tag{3.3.1}$$

e_t 为组合预测在第 t 时刻的预测误差, 则有

$$e_t=x_t-\hat{x}_t=\sum_{i=1}^{m}l_ie_{it}. \tag{3.3.2}$$

设 $J(M)$ 表示组合预测的误差平方和, 则有

$$J(M)=\sum_{t=1}^{N}e_t^2=\sum_{t=1}^{N}\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_je_{it}e_{jt}. \tag{3.3.3}$$

定义 3.3.1 设 $M=\{1, 2,\cdots, m\}$, $s\subset M$, $v(s)$ 为定义在 2^M 集合上实值函数, 令 $v(s)=-J(s)$, 它满足

$$v(\Phi)=0, \tag{3.3.4}$$

$$v(M)\geqslant\sum_{i=1}^{m}v(\{i\}), \tag{3.3.5}$$

则称组合预测方法为一合作 m 人对策, 记为 $\Gamma = [M,\ v]$, $v(s)$ 称为 m 人对策的特征函数, 其中 $J(s)$ 表示联盟 s 进行组合预测所得的预测误差平方和, $v(s)$ 表示 $J(s)$ 的相反数.

因为预测误差平方和越大, 预测精度越低. 所以 $v(s)$ 越大, 表明预测精度越高. 定义 3.3.1 中 (3.3.5) 式表明组合预测方法预测精度高于各单项预测方法精度.

一般来说, 各单项预测方法在组合预测方法中的加权系数的大小应体现其与合作的"贡献"来确定. 下面引进定义 3.3.2.

定义 3.3.2 称 $v(s\cup\{i\})-v(s)$ 为第 i 种各单项预测方法对联盟 s 合作的"贡献", 其中 $s\subset M$.

当所有 m 种单项预测方法均参与组合预测时, $M=\{1,2,\cdots,m\}$为最大的一个联盟, 记 $v(M)$ 为最大的联盟成果, 如何将 $v(M)$ 分配给局中人? 一个自然的想法是依据各局中人给联盟带来的"贡献"来分配.

设 x_i 为第 i 种各单项预测方法从 $v(M)$ 中获得的分配, $i=1,2,\cdots,m$, 则有

$$\begin{aligned}
x_1 &= v(\{1\}),\\
x_2 &= v(\{1,2\})-v(\{1\}),\\
x_3 &= v(\{1,2,3\})-v(\{1,2\}),\\
&\cdots\cdots\cdots\cdots\\
x_m &= v(M)-v(M-\{m\}).
\end{aligned}\tag{3.3.6}$$

然而上述的分配通常与局中人编号的次序有关, 如把局中人 $m,m-1,\cdots,2,1$ 的编号改为 $1',2',\cdots,m'$, 则有新的分配方案

$$\begin{aligned}
x_1' &= v(\{m\}),\\
x_2' &= v(\{m,m-1\})-v(\{m\}),\\
x_3' &= v(\{m,m-1,m-2\})-v(\{m,m-1\}),\\
&\cdots\cdots\cdots\cdots\\
x_m' &= v(M)-v(M-\{1\}).
\end{aligned}\tag{3.3.7}$$

对于其他的编号的次序有对应的分配方案, 由于 m 个局中人编号的次序共有 $m!$ 种, 所以对应的分配方案也有 $m!$ 种, 为此取各局中人分配的平均值作为局中人的平均"贡献".

记 $\varphi_i(v)$ 为第 i 个局中人的平均"贡献", 则有

$$\varphi_i(v)=\frac{1}{m!}\sum_{\pi}\left[v(s_\pi^i\cup\{i\})-v(s_\pi^i)\right],\tag{3.3.8}$$

其中 π 由 $1,2,\cdots,m$ 组成的所有 m 级排列, $\sum$ 为针对所有的 $m!$ 个不同的 m 级排列求和, $s_\pi^i=\{j|\pi j<i\}$.

显然 s_π^i 为排列 π 中排在 i 的前面的那些局中人组成的联盟. 将满足 $s_\pi^i=s$ 排列归为一类, (3.3.8) 式可以表示为

$$\varphi_i(v)=\sum_{i\in s}\frac{(m-|s|)!(|s|-1)!}{m!}\left[v(s)-v(s-\{i\})\right], \quad i=1,2,\cdots,m, \qquad (3.3.9)$$

其中 s 为 M 中包含 $\{i\}$ 的所有子集合, $|s|$ 子集 s 中局中人的单项预测的个数.

定义 3.3.3 称 $\Phi(v)=(\varphi_1(v), \varphi_2(v),\cdots, \varphi_m(v))$ 为合作 m 人对策 $\Gamma=[M, v]$ 的 Shapley 值. 可以证明[57]

$$\sum_{i=1}^{m}\varphi_i(v)=v(M). \qquad (3.3.10)$$

式 (3.3.10) 表明各单项预测方法在组合预测方法中的平均"贡献"之和 $\varphi_i(v)$ 等于合作的总成果.

由 Shapley 值即可计算第 i 种各单项预测方法同联盟合作的平均"贡献" $\varphi_i(v)$. 前已述及, 各单项预测方法在组合预测方法中的加权系数的大小应根据合作中平均"贡献"来确定. 考虑到 $v(M)$ 为预测误差平方和的负值, 需将 $\varphi_i(v)$ 作如下归一化处理, 可得组合预测的加权系数

$$l_i=\frac{v(M)}{\varphi_i(v)}\bigg/\sum_{j=1}^{m}\frac{v(M)}{\varphi_j(v)}, \qquad i=1,2,\cdots,m. \qquad (3.3.11)$$

显然它们满足 $\sum_{i=1}^{m}l_i=1$, $l_i\geqslant 0$, $i=1,2,\cdots,m$.

组合预测权系数确定的合作对策方法计算步骤为

(1) 根据组合预测误差信息矩阵对角线上的元素, 采用某种正权组合方法、如方差倒数加权法、均方差倒数加权法等给出初始的组合预测权系数的估计.

(2) 根据式 (3.3.3) 及定义 3.3.1 计算各种联盟合作的特征函数.

(3) 根据定义 3.3.2 及式 (3.3.9) 计算各种单项预测方法的所获得的平均分配, 即 Sharply 值.

(4) 根据式 (3.3.11) 对各种单项预测方法的所获得的平均分配作归一化处理即得组合预测权系数.

3.3.2 实例分析

设某组合预测问题由 $M=\{1,2,3\}$ 这三种单项预测方法组合而成. 其组合预

测误差信息矩阵为

$$E = (e_{ij})_{3\times 3} = \begin{bmatrix} 10 & 7 & 5 \\ 7 & 8 & 6 \\ 5 & 6 & 6 \end{bmatrix}, \tag{3.3.12}$$

其中 e_{11},e_{22},e_{33} 分别是这三种单项预测方法的预测误差平方和. 这里采用方差倒数加权方法, 其一般计算公式为

$$(l_1, l_2, \cdots, l_m) = \left(\sum_{i=1}^{m} e_{ii}^{-1}\right)^{-1} (e_{11}^{-1}, e_{22}^{-1}, \cdots, e_{mm}^{-1}). \tag{3.3.13}$$

所以按式 (3.3.13) 计算 1, 2, 3 这三种单项预测方法在组合预测方法中的加权系数为

$$(l_1, l_2, l_3) = \left(\frac{1}{10} + \frac{1}{8} + \frac{1}{6}\right)^{-1} \left(\frac{1}{10}, \frac{1}{8}, \frac{1}{6}\right) = \left(\frac{12}{47}, \frac{15}{47}, \frac{20}{47}\right).$$

再按式 (3.3.3) 得

$$J(M) = \left(\frac{12}{47}, \frac{15}{47}, \frac{20}{47}\right) \begin{bmatrix} 10 & 7 & 5 \\ 7 & 8 & 6 \\ 5 & 6 & 6 \end{bmatrix} \left(\frac{12}{47}, \frac{15}{47}, \frac{20}{47}\right)^{\mathrm{T}} = 6.41014,$$

$$v(M) = -J(M) = -6.41014.$$

同理, 采用方差倒数加权方法得

$$v(\{1,2\}) = -7.90123,$$

$$v(\{2,3\}) = -6.3673,$$

$$v(\{1,3\}) = -6.0937,$$

$$v(\{1\}) = -10, \qquad v(\{2\}) = -8, \qquad v(\{3\}) = -6.$$

由式 (3.3.9) 得

$$\begin{aligned} \varphi_1(v) =& v(\{1\}) \times \frac{1}{3} + [v(\{1,2\}) - v(\{2\})] \times \frac{1}{6} + [v(\{1,3\}) - v(\{3\})] \times \frac{1}{6} \\ &+ [v(\{1,2,3\}) - v(\{2,3\})] \times \frac{1}{3} = -3.34676. \end{aligned}$$

同理可得

$$\varphi_2(v) = -2.48356, \qquad \varphi_3(v) = -0.57982.$$

再按式 (3.3.11) 得在组合预测方法中的加权系数为

$$l_1 = 0.123158, \qquad l_2 = 0.165964, \qquad l_3 = 0.710878.$$

$$J'(M) = (0.123158, 0.165964, 0.710878)\begin{bmatrix} 10 & 7 & 5 \\ 7 & 8 & 6 \\ 5 & 6 & 6 \end{bmatrix}\begin{pmatrix} 0.123158 \\ 0.165964 \\ 0.710878 \end{pmatrix}$$
$$= 5.98154.$$

若 1, 2, 3 这三种单项预测方法按以预测误差平方和达到最小的线性组合预测模型计算, 利用 Matlab 最优化工具箱求得在组合预测方法中的最优加权系数为

$$l_1^* = 0.166667, \quad l_2^* = 0, \quad l_3^* = 0.833333.$$

对应的最小预测误差平方和

$$J^*(M) = 5.83333.$$

由此可见, 以预测误差平方和作为预测精度评价准则, 本节提出的确定组合预测方法中的加权系数的 Sharply 值方法明显优于方差倒数加权方法, 且与最优组合预测方法的预测精度很接近.

本节给出了组合预测权系数确定的非最优正权新方法 —— 合作对策方法. 从实例可以看出, 该方法确定的组合预测也是一种优性组合预测[18]. 它不仅适用于以误差平方和为准则的组合预测模型, 而且适用于基于预测有效度的组合预测模型.

3.4 熵值法及其在确定组合预测权系数中的应用

组合预测综合利用各种单项预测方法所提供的信息, 以适当的加权平均形式得出组合预测模型. 组合预测核心的问题就是如何求出组合预测加权平均系数, 使得组合预测模型更加有效地提高预测精度. 那么组合预测加权平均系数是如何依赖于各种单项预测方法所提供的信息以及对信息依赖程度的定量描述等问题值得进一步加强研究. 本节在这方面作一些探索. 即从信息论的观点出发, 根据各单项预测方法预测误差序列的变异程度, 利用信息熵的概念, 计算出组合预测加权平均系数. 实例计算结果令人满意.

3.4.1 确定组合预测加权系数的熵值法的基本原理[58]

设对同一预测对象的某个指标序列为 $\{x_t, t = 1, 2, \cdots, N\}$, 存在 m 种单项预测方法对其进行预测, 设第 i 种单项预测方法在第 t 时刻的预测值为$x_{it}, i =$

$1,2,\cdots,m,t=1,2,\cdots,N$. 令

$$e_{it}=\begin{cases}1, & \text{当}|(x_t-x_{it})/x_t|\geqslant 1\text{时},\\ |(x_t-x_{it})/x_t|, & \text{当}0\leqslant|(x_t-x_{it})/x_t|<1\text{时}.\end{cases} \tag{3.4.1}$$

则称 e_{it} 为第 i 种预测方法第 t 时刻的预测相对误差, $i=1,2,\cdots,m,t=1,2,\cdots,N$. 显然, $0\leqslant e_{it}\leqslant 1,\{e_{it},i=1,2,\cdots,m,t=1,2,\cdots,N\}$ 为第 i 种预测方法在各个 t 时刻的预测相对误差序列.

在信息论中, 熵值是系统无序程度或混乱程度的度量, 信息被解释为系统无序程度的减少, 信息表现为系统的某项指标的变异度. 即系统的熵值越大, 则所蕴涵的信息量越小, 系统的某项指标的变异程度越小. 反之, 系统的熵值越小, 则所蕴涵的信息量越大, 系统的某项指标的变异程度越大.

非最优正权组合预测方法确定组合预测加权系数一个基本思想是: 若某个单项预测模型预测误差序列的变异程度越大, 则其在组合预测中对应的权系数就越小. 在统计学中, 一般采用离差、方差或标准差作为反映预测误差序列变异程度的指标. 然而, 本节利用信息论中熵值的概念, 重新定义单项预测模型预测误差序列的变异系数, 作为反映其变异程度的一个指标.

用熵值法确定组合预测加权系数的步骤如下.

(1) 将各种单项预测方法预测相对误差序列归一化. 即计算第 i 种单项预测方法第 t 时刻的预测相对误差的比重.

$$p_{it}=\frac{e_{it}}{\sum_{t=1}^{N}e_{it}},\qquad t=1,2,\cdots,N. \tag{3.4.2}$$

显然 $\sum_{t=1}^{N}p_{it}=1,i=1,2,\cdots,m$.

(2) 计算第 i 种单项预测方法的预测相对误差的熵值.

$$h_i=-k\sum_{t=1}^{N}p_{it}\ln p_{it},\qquad i=1,2,\cdots,m. \tag{3.4.3}$$

$k>0$ 为常数, ln 为自然对数, $h_i\geqslant 0,i=1,2,\cdots,m,$. 对第 i 种单项预测方法而言, 如果 p_{it} 全部相等, 即 $p_{it}=1/N,t=1,2,\cdots,N$, 那么 h_i 取极大值, 即 $h_i=k\ln N$, 这里取 $k=1/\ln N$, 则有 $0\leqslant h_i\leqslant 1$.

(3) 计算第 i 种单项预测方法的预测相对误差序列的变异程度系数. 因为 $0\leqslant h_i\leqslant 1$, 根据系统某项指标的熵值的大小与其变异程度相反的原则, 所以定义第 i 种

单项预测方法的预测相对误差序列的变异程度系数 d_i 为

$$d_i = 1 - h_i, \quad i = 1, 2, \cdots, m. \tag{3.4.4}$$

(4) 计算各种预测方法的加权系数

$$l_i = \frac{1}{m-1}\left(1 - \frac{d_i}{\sum\limits_{i=1}^{m} d_i}\right), \quad i = 1, 2, \cdots, m. \tag{3.4.5}$$

上式体现了一个原则, 即某个单项预测方法预测误差序列的变异程度越大, 则其在组合预测中对应的权系数就越小. 显然权系数满足 $\sum\limits_{i=1}^{m} l_i = 1$.

(5) 计算组合预测值

$$\hat{x}_t = \sum_{i=1}^{m} l_i x_{it}, \qquad t = 1, 2, \cdots, N. \tag{3.4.6}$$

3.4.2 实例分析

应用文献 [59] 的数据来说明熵值法在确定组合预测权系数中的应用. 已知某预测对象在某一时段的实际观测值 x_t 和两种不同预测方法的预测值 x_{1t}, x_{2t} 见表 3.4.1 所示.

表 3.4.1 两种单项预测方法的预测值

t	x_t	x_{1t}	x_{2t}
1	11.49	18.47	10.03
2	13.06	14.54	11.23
3	15.34	12.84	15.24
4	20.58	13.28	18.67
5	23.28	16.15	27.78
6	26.46	21.16	26.36
7	27.33	28.40	29.67
8	34.22	37.87	27.40
9	40.19	49.58	52.73
10	53.37	63.53	47.36
11	77.79	79.00	71.00
12	100.63	98.12	109.32

按式 (3.4.1) 计算两种不同预测方法的相对误差序列 e_{1t} 和 e_{2t}, 并按式 (3.4.2) 将相对误差序列单位化得 p_{1t} 和 p_{2t}, 结果见表 3.4.2 所示.

由式 (3.4.3) 计算两种单项预测方法的预测相对误差的熵值 h_1, h_2 为

$$h_1 = -\sum_{t=1}^{12} p_{1t} \ln p_{1t} / \ln 12 = 0.8704, \qquad h_2 = -\sum_{t=1}^{12} p_{2t} \ln p_{2t} / \ln 12 = 0.9147.$$

由式 (3.4.4) 计算两种单项预测方法的预测相对误差序列的变异程度系数 d_1, d_2 为

$$d_1 = 1 - h_1 = 0.1296, \quad d_2 = 1 - h_2 = 0.0853.$$

由式 (3.4.5) 计算两种预测方法的加权系数 l_1, l_2 为

$$l_1 = 1 - 0.1296 \ / \ (\ 0.0853 + 0.1296) = 0.3969,$$

$$l_2 = 1 - 0.0853 \ / \ (\ 0.0853 + 0.1296) = 0.6031.$$

所以由式 (3.4.6) 得组合预测值

$$\hat{x}_t = 0.3969 x_{1t} + 0.6031 x_{2t}, \quad t = 1, 2, \cdots, 12.$$

把表 3.4.1 中的数据 x_{1t} 和 x_{2t} 代入上式计算出组合预测值 $\hat{x}_t$, 见表 3.4.2.

表 3.4.2 两种单项预测方法的相对误差序列和归一化

t	e_{1t}	e_{2t}	p_{1t}	p_{2t}
1	0.6075	0.1271	0.2584	0.1061
2	0.1133	0.1401	0.0482	0.1170
3	0.1630	0.0065	0.0693	0.0054
4	0.3499	0.0928	0.1488	0.0775
5	0.3063	0.1933	0.1303	0.1614
6	0.2003	0.0038	0.0852	0.0032
7	0.0392	0.0856	0.0167	0.0715
8	0.1067	0.1993	0.0454	0.1664
9	0.2336	0.0632	0.0994	0.0528
10	0.1904	0.1126	0.0810	0.0940
11	0.0156	0.0873	0.0066	0.0729
12	0.0249	0.0864	0.0106	0.0721

若以预测误差平方和作为反映预测精度的一个指标, 则有熵值法确定的组合预测模型对应的预测误差平方和为

$$Q_1 = \sum_{t=1}^{12} e_t^2 = \sum_{t=1}^{12} (x_t - \hat{x}_t)^2 = 95.0940.$$

文献 [38] 已给出以预测误差平方和为准则的非负线性组合预测模型的计算最优组合预测权系数 l_1^*, l_2^* 得

$$l_1^* = 0.4121, \qquad l_2^* = 0.5879.$$

模型 (3.1.2) 的目标函数最优值 Q_2 就是预测误差平方和, 则

$$Q_2 = 94.8889,$$

$$Q_1 = \sum_{t=1}^{12} e_t^2 = \sum_{t=1}^{12} (x_t - \hat{x}_t)^2 = 95.0940.$$

由此可见, 以预测误差平方和作为预测精度评价准则, 本节提出的确定组合预测方法中的加权系数的熵值方法与最优组合预测方法的预测精度很接近. 另外, 通过比较熵值方法和方差倒数加权方法, 也发现本节提出的熵值方法稍优于方差倒数加权方法. 这表明熵值方法用于确定组合预测方法加权系数具有一定的有效性. 从本节实例可以看出, 该方法确定的组合预测也是一种优性组合预测 [18].

第 4 章　基于预测误差指标的最优组合预测模型

4.1　预 备 知 识

数学规划研究函数在各种条件下的极值问题. 在最优组合预测模型的研究中, 数学规划模型和求解方法有着重要的应用. 因此本节介绍这些方面的相关知识.

4.1.1　凸集和凸函数[60]

定义 4.1.1　设 S 是 R^n 的子集, 若对于任意的 $\boldsymbol{s}_1, \boldsymbol{s}_2 \in S$, 均有

$$\lambda \boldsymbol{s}_1 + (1-\lambda)\boldsymbol{s}_2 \in S, \quad \forall \lambda \in [0,1], \tag{4.1.1}$$

则称 S 是 R^n 的凸集.

定义 4.1.2　设函数 $f: S \to R^1$, S 是 R^n 的非空凸子集, 若

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leqslant \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2), \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S, \lambda \in (0,1), \tag{4.1.2}$$

则称 f 为 S 上的凸函数, 当上式成立严格不等号时, 则称 f 为 S 上的严格凸函数.

定义 4.1.3　设函数 $f: S \to R^1$ 具有二阶连续的偏导数, $\boldsymbol{x}^* \in S$, 令

$$\nabla f(\boldsymbol{x}^*) = \left(\frac{\partial f(\boldsymbol{x}^*)}{\partial x_1}, \frac{\partial f(\boldsymbol{x}^*)}{\partial x_2}, \cdots, \frac{\partial f(\boldsymbol{x}^*)}{\partial x_n}\right)^{\mathrm{T}}, \quad H(\boldsymbol{x}^*) = \left(\frac{\partial^2 f(\boldsymbol{x}^*)}{\partial x_i \partial x_j}\right)_{n\times n},$$

则称 $\nabla f(\boldsymbol{x}^*)$ 为函数 f 在点 $\boldsymbol{x}^*$ 处的梯度向量, 称 $H(\boldsymbol{x}^*)$ 为函数 f 在点 $\boldsymbol{x}^*$ 处的 Hesse 矩阵.

定理 4.1.1　设 S 是 R^n 的非空开凸集, 函数 $f: S \to R^1$ 具有一阶连续的偏导数, 则函数 f 为凸函数的充要条件是恒有

$$f(\boldsymbol{x}_2) \geqslant f(\boldsymbol{x}_1) + \nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2 - \boldsymbol{x}_1), \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S, \tag{4.1.3}$$

其中 $\nabla f(\boldsymbol{x}_1)$ 为函数 f 在 $\boldsymbol{x}_1$ 处的梯度向量.

证明　必要性. $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, 因为 S 是 R^n 的非空开凸集, 所以

$$\boldsymbol{x}_1 + \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1) = (1-\lambda)\boldsymbol{x}_1 + \lambda \boldsymbol{x}_2 \in S, \quad \forall \lambda \in [0,1].$$

因为函数 f 为凸函数, 所以

$$f(\boldsymbol{x}_1 + \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1)) \leqslant \lambda f(\boldsymbol{x}_2) + (1-\lambda) f(\boldsymbol{x}_1).$$

从而有

$$f\left(\boldsymbol{x}_1+\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1)\right)-f(\boldsymbol{x}_1)\leqslant \lambda\left(f(\boldsymbol{x}_2)-f(\boldsymbol{x}_1)\right).$$

由于 $\lambda>0$, 上式不等式两边同除以 λ 得

$$\frac{f\left(\boldsymbol{x}_1+\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1)\right)-f(\boldsymbol{x}_1)}{\lambda}\leqslant f(\boldsymbol{x}_2)-f(\boldsymbol{x}_1).$$

由于 f 具有一阶连续的偏导数, 故有

$$f\left(\boldsymbol{x}_1+\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1)\right)=f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}\left[\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1)\right]+o\left(\|\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1)\|\right).$$

所以

$$\lim_{\lambda\to 0+}\frac{f\left(\boldsymbol{x}_1+\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1)\right)-f(\boldsymbol{x}_1)}{\lambda}=\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1).$$

于是得到

$$f(\boldsymbol{x}_2)\geqslant f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1).$$

充分性. $\forall \boldsymbol{x}_1,\boldsymbol{x}_2\in S$, $\lambda\in(0,1)$, 因为 S 是 R^n 的非空开凸集, 则 $\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in S$, 由充分性条件式 (4.1.3) 得

$$f(\boldsymbol{x}_1)\geqslant f\left(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\right)+\nabla f(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2)^{\mathrm{T}}\left[(1-\lambda)(\boldsymbol{x}_1-\boldsymbol{x}_2)\right],$$

$$f(\boldsymbol{x}_2)\geqslant f\left(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\right)+\nabla f(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2)^{\mathrm{T}}\left[-\lambda(\boldsymbol{x}_1-\boldsymbol{x}_2)\right].$$

以 λ 和 $1-\lambda$ 分别乘以上述两个不等式的两边, 并相加整理得

$$\lambda f(x_1)+(1-\lambda)f(\boldsymbol{x}_2)\geqslant f\left(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\right),\quad \forall \boldsymbol{x}_1,\boldsymbol{x}_2\in S,\lambda\in(0,1).$$

由定义 4.1.2 知函数 f 为凸函数.

定理 4.1.2　设 S 是 R^n 的非空开凸集, 函数 $f:S\to R^1$ 具有二阶连续的偏导数, 函数 f 为凸函数的充要条件是函数 f 的 Hesse 矩阵 $H(\boldsymbol{x})$ 在 S 上为半正定矩阵. 函数 f 为严格凸函数的充要条件是函数 f 的 Hesse 矩阵 $H(\boldsymbol{x})$ 在 S 上为正定矩阵.

证明　必要性. 因为 S 是 R^n 的非空开凸集, 所以

$$\forall\lambda\in[0,1],\quad \boldsymbol{x}_1+\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1)=(1-\lambda)\boldsymbol{x}_1+\lambda\boldsymbol{x}_2\in S.$$

函数 f 为严格凸函数, $\forall \boldsymbol{x}\in S$, $\boldsymbol{z}\in R^n$, 因为 S 是 R^n 的非空开凸集, 所以存在 $\bar\lambda>0$, 当 $\lambda\in[-\bar\lambda,\bar\lambda]$ 时, 有 $\boldsymbol{x}+\lambda\boldsymbol{z}\in S$, 由定理 4.1.1 得

$$f\left(\boldsymbol{x}+\lambda\boldsymbol{z}\right)\geqslant f(\boldsymbol{x})+\lambda\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{z}.$$

因为函数 $f: S\to R^1$ 具有二阶连续的偏导数, 由 Taylor 展式得

$$f(\boldsymbol{x}+\lambda\boldsymbol{z})=f(\boldsymbol{x})+\lambda\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{z}+\frac{1}{2}\lambda^2\boldsymbol{z}^{\mathrm{T}}H(\boldsymbol{x})\boldsymbol{z}+o(\lambda^2).$$

所以由上两式得

$$\frac{1}{2}\lambda^2\boldsymbol{z}^{\mathrm{T}}H(\boldsymbol{x})\boldsymbol{z}+o(\lambda^2)\geqslant 0.$$

上式两边同除以 λ^2, 注意到 $\lim\limits_{\lambda\to 0}\dfrac{o(\lambda^2)}{\lambda^2}=0$, 则有

$$\boldsymbol{z}^{\mathrm{T}}H(\boldsymbol{x})\boldsymbol{z}\geqslant 0.$$

即 Hesse 矩阵 $H(\boldsymbol{x})$ 在 S 上为半正定矩阵.

充分性. $\forall\bar{\boldsymbol{x}}\in S$, 由 Taylor 展式得

$$f(\boldsymbol{x})=f(\bar{\boldsymbol{x}})+\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})+\frac{1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}H(\bar{\boldsymbol{x}}+\theta(\boldsymbol{x}-\bar{\boldsymbol{x}}))(\boldsymbol{x}-\bar{\boldsymbol{x}}),\quad \forall\boldsymbol{x},$$

其中 $\theta\in(0,1)$. 因为 S 是凸集, 所以 $\bar{\boldsymbol{x}}+\theta(\boldsymbol{x}-\bar{\boldsymbol{x}})\in S$, 由充分性的条件知 Hesse 矩阵 $H(\boldsymbol{x})$ 在 S 上为半正定矩阵, 从而 $H(\bar{\boldsymbol{x}}+\theta(\boldsymbol{x}-\bar{\boldsymbol{x}}))$ 为半正定矩阵, 即有 $(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}H(\bar{\boldsymbol{x}}+\theta(\boldsymbol{x}-\bar{\boldsymbol{x}}))(\boldsymbol{x}-\bar{\boldsymbol{x}})\geqslant 0$, 所以

$$f(\boldsymbol{x})\geqslant f(\bar{\boldsymbol{x}})+\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\quad \forall\boldsymbol{x}.$$

由定理 4.1.1 知函数 f 为 S 的凸函数.

第二个结论的证明只要把上述的不等号换成严格的不等号即知成立.

4.1.2 非线性规划问题的最优性条件 [60~62]

下面考虑如下非线性规划问题的最优性条件.

$$\begin{aligned}&\min f(\boldsymbol{x}),\\ &\text{s.t.}\begin{cases} g_i(\boldsymbol{x})\leqslant 0,\\ i=1,2,\cdots,m.\end{cases}\end{aligned}\tag{4.1.4}$$

定义 4.1.4 记 $S=\{\boldsymbol{x}\,|g_i(\boldsymbol{x})\leqslant 0, i=1,2,\cdots,m\}$, 则称 S 为式 (4.1.4) 表示的非线性规划问题的可行域, $\forall\boldsymbol{x}\in S$ 为其可行解.

定义 4.1.5 记 $N=\{1,2,\cdots,m\}$, 若对某 $i\in N$, $\boldsymbol{x}^*\in S$, 有 $g_i(\boldsymbol{x}^*)=0$, 则称 $g_i(\boldsymbol{x})\leqslant 0$ 为关于 $\boldsymbol{x}^*$ 的积极约束. 称 $N(\boldsymbol{x}^*)=\{i|\,g_i(\boldsymbol{x}^*)=0, i\in N\}$ 为积极约束指标集合, 若在 $\boldsymbol{x}^*$ 处, 梯度向量 $\nabla g_i(\boldsymbol{x}^*)\,(\forall i\in N(\boldsymbol{x}^*))$ 线性无关, 则称 $\boldsymbol{x}^*$ 为正则点.

记

$$F(\boldsymbol{x}^*)=\left\{\boldsymbol{p}|\,\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{p}<0\right\},\quad G(\boldsymbol{x}^*)=\bigcap_{i\in N(\boldsymbol{x}^*)}\left\{\boldsymbol{p}|\,\nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{p}<0\right\}.$$

定理 4.1.3　设函数 $f: R^n \to R^1$ 和 $g_i: R^n \to R^1 (i \in N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处可微, $g_i(i \notin N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处连续, 若 $\boldsymbol{x}^*$ 是非线性规划问题式 (4.1.4) 的局部最优解, 则有

$$F(\boldsymbol{x}^*) \cap G(\boldsymbol{x}^*) = \varnothing.$$

证明　采用反证法. 假设 $F(\boldsymbol{x}^*) \cap G(\boldsymbol{x}^*) \neq \emptyset$, 则存在 $\boldsymbol{p} \in F(\boldsymbol{x}^*) \cap G(\boldsymbol{x}^*)$, 由 $F(\boldsymbol{x}^*)$ 和 $G(\boldsymbol{x}^*)$ 的意义可知

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{p} < 0, \quad \nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{p} < 0, \quad i \in N(\boldsymbol{x}^*). \tag{4.1.5}$$

因为函数 $f: R^n \to R^1$ 和 $g_i: R^n \to R^1 (i \in N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处可微, 在 $\boldsymbol{x}^*$ 处对它们进行 Taylor 展开, $\forall t > 0$, 有

$$f(\boldsymbol{x}^* + t\boldsymbol{p}) = f(\boldsymbol{x}^*) + t\nabla f(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{p} + o\,\|t\boldsymbol{p}\|, \tag{4.1.6}$$

$$g_i(\boldsymbol{x}^* + t\boldsymbol{p}) = g_i(\boldsymbol{x}^*) + t\nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{p} + o\,\|t\boldsymbol{p}\|, \quad i \in N(\boldsymbol{x}^*). \tag{4.1.7}$$

当 $t > 0$ 充分小时, 由式 (4.1.5)~(4.1.7) 知

$$f(\boldsymbol{x}^* + t\boldsymbol{p}) < f(\boldsymbol{x}^*), \quad g_i(\boldsymbol{x}^* + t\boldsymbol{p}) < g_i(\boldsymbol{x}^*) = 0, \quad i \in N(\boldsymbol{x}^*).$$

因为 $g_i(i \notin N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处连续, 所以当 $t(>0)$ 充分小时, $g_i(\boldsymbol{x}^*+t\boldsymbol{p}) \leqslant 0, i \notin N(\boldsymbol{x}^*)$.

综上所述, 找到了与 $\boldsymbol{x}^*$ 处充分接近的可行解 $\boldsymbol{x}^* + t\boldsymbol{p}$, 使得 $f(\boldsymbol{x}^* + t\boldsymbol{p}) < f(\boldsymbol{x}^*)$, 这与 $\boldsymbol{x}^*$ 是式 (4.1.4) 的局部最优解矛盾, 因此假设不成立, 从而结论成立.

定理 4.1.4　设函数 $f: R^n \to R^1$ 和 $g_i: R^n \to R^1 (i \in N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处可微, $g_i(i \notin N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处连续, 若 $\boldsymbol{x}^*$ 是非线性规划问题式 (4.1.4) 的局部最优解, 则存在实数 μ_0, $(\mu_i, \forall i \in N(\boldsymbol{x}^*))$ 满足

$$\mu_0 \nabla f(\boldsymbol{x}^*) + \sum_{i \in N(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}, \tag{4.1.8}$$

$$\mu_0 \geqslant 0, \quad \mu_i \geqslant 0, \quad (\mu_0, \mu_i) \neq \boldsymbol{0}, \quad \forall i \in N(\boldsymbol{x}^*). \tag{4.1.9}$$

式 (4.1.8)、(4.1.9) 称为 $\boldsymbol{x}^*$ 满足 Fritz-John 条件.

证明　由定理 4.1.3 知, 不存在 $\boldsymbol{p} \in F(\boldsymbol{x}^*) \cap G(\boldsymbol{x}^*)$, 即不等式组

$$\begin{cases} \nabla f(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{p} < 0, \\ \nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{p} < 0, \quad i \in N(\boldsymbol{x}^*) \end{cases}$$

不存在解 $\boldsymbol{p}$, 令矩阵 $\boldsymbol{A} = \begin{bmatrix} \nabla f(\boldsymbol{x}^*)^{\mathrm{T}} \\ \left(\nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}\right)_{i \in N(\boldsymbol{x}^*)} \end{bmatrix}$, 其中 $\left(\nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}\right)_{i \in N(\boldsymbol{x}^*)}$ 是以 $\nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}$ 为行的矩阵, 于是 $\boldsymbol{A}\boldsymbol{p} < \boldsymbol{0}$ 无解. 由 Gordan 择一性定理, 系统 $\boldsymbol{A}^{\mathrm{T}} \boldsymbol{p} = \boldsymbol{0}$,

$\boldsymbol{p} \geqslant \boldsymbol{0}, \boldsymbol{p} \neq \boldsymbol{0}$ 必有解, 记其解 $\boldsymbol{p} = (\mu_0, (\mu_i)_{i \in N(\boldsymbol{x}^*)})$, 则 $\mu_0 \geqslant 0$, $\mu_i \geqslant 0$, $\forall i \in N(\boldsymbol{x}^*)$ 但 $(\mu_i, \forall i \in N(\boldsymbol{x}^*)) \neq \boldsymbol{0}$, $\mu_0 \geqslant 0$, 由 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{p} = \boldsymbol{0}$ 展开即得

$$\mu_0 \nabla f(\boldsymbol{x}^*) + \sum_{i \in N(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}.$$

若对所有的 $g_i : R^n \to R^1 (i \in N)$ 在 $\boldsymbol{x}^*$ 处可微, 则 Fritz-John 条件可表述为: 存在不全为 0 的实数 $\mu_0, \mu_1, \mu_2, \cdots, \mu_m$, 满足

$$\mu_0 \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}, \quad \mu_i g_i(\boldsymbol{x}^*) = 0, \quad i = 1, 2, \cdots, m,$$

$$(\mu_0, \mu_1, \mu_2, \cdots, \mu_m) \geqslant \boldsymbol{0}, \quad (\mu_0, \mu_1, \mu_2, \cdots, \mu_m) \neq \boldsymbol{0}.$$

定义 4.1.6 设 $\boldsymbol{x}^* \in S$ 为非线性规划问题式 (4.1.4) 的可行解, 若存在标量 $(\mu_i, \forall i \in N(\boldsymbol{x}^*))$ 满足

$$\nabla f(\boldsymbol{x}^*) + \sum_{i \in N(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}, \quad \mu_i \geqslant 0, \quad \forall i \in N(\boldsymbol{x}^*). \tag{4.1.10}$$

则称在 $\boldsymbol{x}^*$ 为 Kuhn-Tucker 点或 $\boldsymbol{x}^*$ 满足 Kuhn-Tucker 条件 (4.1.10).

若对所有的 $g_i : R^n \to R^1 (i = 1, 2, \cdots, m)$ 在 $\boldsymbol{x}^*$ 处可微, 则 Kuhn-Tucker 可表述为: 存在非负实数 $(\mu_1, \mu_2, \cdots, \mu_m) \geqslant \boldsymbol{0}$, 满足

$$\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}, \quad \mu_i g_i(\boldsymbol{x}^*) = 0, \quad i = 1, 2, \cdots, m.$$

定理 4.1.5 设函数 $f : R^n \to R^1$ 和 $g_i : R^n \to R^1 (i \in N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处可微, $g_i (i \notin N(\boldsymbol{x}^*))$ 在 $\boldsymbol{x}^*$ 处连续, 若 $\boldsymbol{x}^*$ 是非线性规划问题式 (4.1.4) 的局部最优解, $\boldsymbol{x}^*$ 是正则点, 则 $\boldsymbol{x}^*$ 为 Kuhn-Tucker 点.

证明 由定理 4.1.4 知, 存在满足 Fritz-John 条件的实数 μ_0, $(\mu_i, \forall i \in N(\boldsymbol{x}^*))$ 满足

$$\begin{gathered} \mu_0 \nabla f(\boldsymbol{x}^*) + \sum_{i \in N(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}, \\ \mu_0 \geqslant 0, \quad \mu_i \geqslant 0, \quad \forall i \in N(\boldsymbol{x}^*), \quad (\mu_0, \mu_i, \forall i \in N(\boldsymbol{x}^*)) \neq \boldsymbol{0}. \end{gathered} \tag{4.1.11}$$

若 $\mu_0 = 0$, 则有 $\displaystyle\sum_{i \in N(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}$, 且

$$\mu_i \geqslant 0, \forall i \in N(\boldsymbol{x}^*), \quad (\mu_i, \forall i \in N(\boldsymbol{x}^*)) \neq \boldsymbol{0},$$

此时 $\nabla g_i(\boldsymbol{x}^*)$, $\forall i \in N(\boldsymbol{x}^*)$ 线性相关, 这与 $\boldsymbol{x}^*$ 是正则点矛盾, 所以 $\mu_0 \neq 0$, 从而 $\mu_0 > 0$, 令 $\nu_i = \dfrac{\mu_i}{\mu_0}, \forall i \in N(\boldsymbol{x}^*)$, (4.1.11) 两边同除以 μ_0 得

$$\nabla f(\boldsymbol{x}^*) + \sum_{i \in N(\boldsymbol{x}^*)} \nu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}, \quad \nu_i \geqslant 0, \quad \forall i \in N(\boldsymbol{x}^*).$$

由定义 4.1.6 知结论成立.

定理 4.1.6 设函数 $f: R^n \to R^1$ 和 $g_i: R^n \to R^1, i \in N(\boldsymbol{x}^*)$ 在 $\boldsymbol{x}^*$ 处可微, g_i, $i \notin N(\boldsymbol{x}^*)$ 在 $\boldsymbol{x}^*$ 处连续, 函数 f 和 g_i, $i \in N(\boldsymbol{x}^*)$ 是凸函数, 若 $\boldsymbol{x}^*$ 为 Kuhn-Tucker 点, 则 $\boldsymbol{x}^*$ 是非线性规划问题式 (4.1.4) 的最优解.

证明 设 $\forall \boldsymbol{x} \in S$ 是非线性规划问题式 (4.1.4) 的可行解, 则有 $g_i(\boldsymbol{x}) \leqslant 0 = g_i(\boldsymbol{x}^*), \forall i \in N(\boldsymbol{x}^*)$. 因为 g_i, $i \in N(\boldsymbol{x}^*)$ 是凸函数, 所以有

$$\begin{aligned} g_i(\boldsymbol{x}^* + \lambda(\boldsymbol{x} - \boldsymbol{x}^*)) &= g_i(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}^*) \\ &\leqslant \lambda g_i(\boldsymbol{x}) + (1-\lambda) g_i(\boldsymbol{x}^*) \leqslant g_i(\boldsymbol{x}^*), \quad \forall \lambda \in (0,1). \end{aligned} \tag{4.1.12}$$

因为 g_i, $i \in N(\boldsymbol{x}^*)$ 在 $\boldsymbol{x}^*$ 处可微, 对它们进行 Taylor 展开得

$$g_i(\boldsymbol{x}^* + \lambda(\boldsymbol{x} - \boldsymbol{x}^*)) = g_i(\boldsymbol{x}^*) + \lambda \nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) + o\left\|\lambda(\boldsymbol{x} - \boldsymbol{x}^*)\right\|. \tag{4.1.13}$$

由式 (4.1.12) 和式 (4.1.13) 知

$$\nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \leqslant 0, \quad \forall i \in N(\boldsymbol{x}^*).$$

从而对于 $\mu_i \geqslant 0, i \in N(\boldsymbol{x}^*)$ 有

$$\sum_{i \in N(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \leqslant 0.$$

因为 $\boldsymbol{x}^*$ 为 Kuhn-Tucker 点, 所以有 $\nabla f(\boldsymbol{x}^*) + \sum\limits_{i \in N(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}$, 于是有

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \geqslant 0, \quad \forall \boldsymbol{x} \in S.$$

因为函数 f 是凸函数, 所以 $f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*)$, $\forall \boldsymbol{x} \in S$, 即 $\boldsymbol{x}^*$ 是非线性规划问题式 (4.1.4) 的最优解.

4.2 以预测误差平方和达到最小的线性组合预测模型

4.2.1 最优线性组合预测模型的建立

设对同一预测对象的某个指标序列 $\{x_1, t = 1, 2, \cdots, N\}$, 存在 m 种单项无偏预测方法对其进行预测, 设第 i 种单项预测方法在第 t 时刻的预测值为 x_{it}, $i =$

$1,2,\cdots,m$, $t=1,2,\cdots,N$, 称 $e_{it}=(x_t-x_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差.

设 $l_1,l_2,\cdots,l_m$ 分别为 m 种单项预测方法的加权系数, 为了使组合预测保持无偏性, 加权系数应满足

$$l_1+l_2+\cdots+l_m=1, \tag{4.2.1}$$

设 $\hat{x}_t=l_1x_{1t}+l_2x_{2t}+\cdots+l_mx_{mt}$ 为 x_t 的组合预测值, 设 e_t 为组合预测在第 t 时刻的预测误差, 则有

$$e_t=x_t-\hat{x}_t=\sum_{i=1}^{m}l_ie_{it}. \tag{4.2.2}$$

设 J_1 表示组合预测预测误差平方和, 则有

$$J_1=\sum_{t=1}^{N}e_t^2=\sum_{t=1}^{N}\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_je_{it}e_{jt}. \tag{4.2.3}$$

由此可得以预测误差平方和为准则的线性组合预测模型为下列最优化问题

$$\min J_1=\sum_{t=1}^{N}\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_je_{it}e_{jt},$$
$$\text{s.t.}\sum_{i=1}^{m}l_i=1. \tag{4.2.4}$$

记

$$\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}},\quad \boldsymbol{R}=(1,1,\cdots,1)^{\mathrm{T}},\quad \boldsymbol{e}_i=(e_{i1},e_{i2},\cdots,e_{iN})^{\mathrm{T}},$$

则 $\boldsymbol{L}$ 表示组合预测加权系数列向量, $\boldsymbol{R}$ 表示元素全为 1 的 m 维列向量, $\boldsymbol{e}_i$ 表示第 i 种单项预测方法的预测误差列向量, 再令

$$E_{ij}=\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{e}_j=\sum_{t=1}^{N}e_{it}e_{jt},\quad i,j=1,2,\cdots,m,\quad \boldsymbol{E}=(E_{ij})_{m\times m},$$

则当 $i\neq j$ 时, E_{ij} 表示第 i 种单项预测方法和第 j 种单项预测方法的预测误差的协方差, 当 $i=j$ 时, E_{ii} 表示第 i 种单项预测方法的预测误差的平方和, $\boldsymbol{E}$ 表示 $m\times m$ 的方阵, $\boldsymbol{E}$ 称为组合预测误差信息矩阵.

定理 4.2.1 假定 $m(m<N)$ 种单项预测方法的预测误差向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 则组合预测误差信息矩阵 $\boldsymbol{E}$ 为正定矩阵.

证明　记 $\boldsymbol{A}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m)$, 则

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}\boldsymbol{e}_1^{\mathrm{T}}\\ \boldsymbol{e}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_m^{\mathrm{T}}\end{pmatrix}(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m)=\begin{pmatrix}\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_1 & \boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_2 & \cdots & \boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_m\\ \boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_1 & \boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2 & \cdots & \boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_m\\ \vdots & \vdots & & \vdots\\ \boldsymbol{e}_m^{\mathrm{T}}\boldsymbol{e}_1 & \boldsymbol{e}_m^{\mathrm{T}}\boldsymbol{e}_2 & \cdots & \boldsymbol{e}_m^{\mathrm{T}}\boldsymbol{e}_m\end{pmatrix}=\boldsymbol{E}.$$

所以 $\boldsymbol{E}$ 为对称矩阵, 记 $\boldsymbol{x}=(x_1,x_2,\cdots,x_m)^{\mathrm{T}}$, 若 $\boldsymbol{x}\neq\boldsymbol{0}$ 时, 则 $\boldsymbol{A}\boldsymbol{x}\neq\boldsymbol{0}$, 否则, 存在不全为零的数 $x_1,x_2,\cdots,x_m$, 使得

$$\boldsymbol{A}\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_m\boldsymbol{e}_m=\boldsymbol{0}.$$

此与向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的矛盾!

当 $\boldsymbol{x}\neq\boldsymbol{0}$ 时, 二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$, 所以二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}$ 为正定二次型. 即组合预测误差信息矩阵 $\boldsymbol{E}$ 为正定矩阵.

定理 4.2.1 表明在预测误差向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的条件下, 组合预测误差信息矩阵 $\boldsymbol{E}$ 为可逆矩阵.

在上述记号下, 有

$$\begin{aligned}J_1&=\sum_{t=1}^{N}\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_je_{it}e_{jt}=\sum_{i=1}^{m}\sum_{j=1}^{m}\left[l_il_j\left(\sum_{t=1}^{N}e_{it}e_{jt}\right)\right]\\&=\sum_{i=1}^{m}\sum_{j=1}^{m}[l_il_jE_{ij}]=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}.\end{aligned}\tag{4.2.5}$$

$$\sum_{i=1}^{m}l_i=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}.\tag{4.2.6}$$

所以式 (4.2.4) 也可以表示成矩阵形式

$$\begin{aligned}&\min J_1=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L},\\&\text{s.t.}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}=1.\end{aligned}\tag{4.2.7}$$

4.2.2 最优线性组合预测模型的解的讨论

对于上式 (4.2.7), 可给出如下组合预测加权系数的计算公式.

定理 4.2.2[14]　假定 $m(m<N)$ 种单项预测方法的预测误差向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 则有模型 (4.2.7) 的最优解和目标函数最优值为

$$\boldsymbol{L}^*=\frac{\boldsymbol{E}^{-1}\boldsymbol{R}}{\boldsymbol{R}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{R}},\quad J_1^*=\frac{1}{\boldsymbol{R}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{R}}.\tag{4.2.8}$$

证明 利用 Lagrange 乘数法易得, 证略.

在实际预测实践中, 若利用 (4.2.8) 来计算组合预测加权系数, 则可能出现加权系数为负的情况. 而负的组合预测加权系数的解释在学术界尚未取得一致意见, 因此有必要考虑非负权系数的组合预测模型. 这就要在模型 (4.2.7) 中增加一个非负约束条件, 即为如下最优化模型

$$\min J_2 = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}, \quad \text{s.t.} \begin{cases} \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1, \\ \boldsymbol{L} \geqslant \boldsymbol{0}, \end{cases} \tag{4.2.9}$$

上式 (4.2.9) 为一个非线性规划问题. 对于该问题有如下结论.

定理 4.2.3 假定 $m(m < N)$ 种单项预测方法的预测误差向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_m$ 是线性无关的, 则非负权系数的组合预测模型 (4.2.9) 的可行域 $D = \{\boldsymbol{L}|\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1, \boldsymbol{L} \geqslant \boldsymbol{0}\}$ 为凸集, 其目标函数 $J_2 = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}$ 为可行域 D 上的严格凸函数.

证明 设 $\forall \boldsymbol{L}_1, \boldsymbol{L}_2 \in D$, 由可行域 D 的定义知

$$\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_1 = 1, \quad \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_2 = 1, \quad \boldsymbol{L}_1 \geqslant \boldsymbol{0}, \quad \boldsymbol{L}_2 \geqslant \boldsymbol{0}.$$

所以

$$\boldsymbol{R}^{\mathrm{T}}(\lambda\boldsymbol{L}_1 + (1-\lambda)\boldsymbol{L}_2) = \lambda\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_1 + (1-\lambda)\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_2 = \lambda + (1-\lambda) = 1, \quad \forall\lambda \in [0,1].$$

且 $\lambda\boldsymbol{L}_1 + (1-\lambda)\boldsymbol{L}_2 \geqslant \boldsymbol{0}$, 所以由定义 (4.1.1) 知可行域 D 为凸集.

由于 m 种单项预测方法的预测误差向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_m$ 是线性无关的, 由定理 4.2.1 知组合预测误差信息矩阵 $\boldsymbol{E}$ 为正定矩阵, 且模型 (4.2.9) 的目标函数 $J_2 = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}$ 的 Hesse 矩阵

$$H(\boldsymbol{L}) = \left(\frac{\partial^2 J_2}{\partial l_i \partial l_j}\right)_{m\times m} = \boldsymbol{E}.$$

再根据定理 4.1.2 得目标函数 J_2 是可行域 D 上的严格凸函数.

模型 (4.2.9) 实际上为一个二次凸规划问题. 由定理 4.1.6 可知 Kuhn-Tucker 条件是其最优解的充要条件. 模型 (4.2.9) 的 Kuhn-Tucker 条件可表示为

$$\begin{cases} 2\boldsymbol{E}\boldsymbol{L} - \lambda\boldsymbol{R} - \boldsymbol{U} = \boldsymbol{0}, \\ \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1, \\ \boldsymbol{U}^{\mathrm{T}}\boldsymbol{L} = 0, \\ \boldsymbol{L} \geqslant \boldsymbol{0}, \quad \boldsymbol{U} \geqslant \boldsymbol{0}, \end{cases} \tag{4.2.10}$$

其中 $\boldsymbol{U}=(u_1,u_2,\cdots,u_m)^{\mathrm{T}}$ 是与非负组合预测权系数向量 $\boldsymbol{L}$ 所对应的 Kuhn-Tucker 乘子, u_i 与 l_i 不能同时为基变量. λ 是与约束条件 $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}=1$ 所对应的 Lagrange 乘子.

由于 λ 无非负约束, 可令 $\lambda=\lambda_1-\lambda_2$, 其中 $\lambda_1,\lambda_2\geqslant 0$, 因此, 引入人工变量 v 可构造如下线性规划模型

$$\min v,$$
$$\text{s.t.}\begin{cases}\boldsymbol{2EL}-(\lambda_1-\lambda_2)\boldsymbol{R}-\boldsymbol{U}=\boldsymbol{0},\\ \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}+v=1,\\ \boldsymbol{L}\geqslant\boldsymbol{0},\boldsymbol{U}\geqslant\boldsymbol{0},\\ \lambda_1,\lambda_2,\quad v\geqslant 0.\end{cases}\tag{4.2.11}$$

解此线性规划模型即可获得最优的非负组合预测权系数向量.

4.3 以误差绝对值和达到最小的线性组合预测模型

众所周知, 预测误差平方和是反映预测精度的一个非常重要的指标之一. 因此, 以预测误差平方和达到最小为准则的线性组合预测模型目前在各领域实际预测问题中应用得最为广泛. 造成这种现象的主要原因可能包括有两方面: 一是以预测误差平方和达到最小为准则的线性组合预测模型理论上研究比较深入; 二是通过前面的定理 4.2.2 可知, 对于组合预测权系数无非负约束的最优组合预测模型, 它的最优解有明显的数学表达式, 从而可套公式计算.

然而, 预测误差平方和作为预测精度的指标存在一定的缺陷. 这主要是预测误差平方和受异常点数据影响较大. 也就是说, 在数据中含有异常数据, 这使得预测误差在该点较大, 预测误差再平方后就会使该点预测误差就更大了, 即预测误差平方和在数据异常点处产生预测误差 "放大" 效应. 在处理异常点数据时, 若凭经验给予剔除, 则可能造成一些信息的损失. 因为异常点数据在某些方面确实包含一些特殊有用的信息. 考虑了上述预测误差平方和的缺陷, 有必要引进更加稳健的指标来刻画预测精度.

所谓稳健性是指当统计数据发生较小的变化时, 由组合预测模型计算的组合预测权系数也发生较小的变化. 于是, 人们给出另外一个非常重要的指标, 即误差绝对值和. 它的稳健性比预测误差平方和要好得多.

对于线性组合预测 $\hat{x}_t=l_1x_{1t}+l_2x_{2t}+\cdots+l_mx_{mt}$, 设 e_t 为组合预测在第 t 时刻的预测误差, 则有

$$e_t=x_t-\hat{x}_t=\sum_{i=1}^{m}l_ie_{it},$$

其中 $e_{it} = (x_t - x_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差, $l_1, l_2, \cdots, l_m$ 分别为 m 种单项预测方法的加权系数, 且满足

$$l_1 + l_2 + \cdots + l_m = 1, \quad l_1, l_2, \cdots, l_m \geqslant 0.$$

设 F 表示组合预测误差绝对值和, 则有

$$F = \sum_{t=1}^{N} |e_t| = \sum_{t=1}^{N} \left| \sum_{i=1}^{m} l_i e_{it} \right|. \tag{4.3.1}$$

可得以预测误差绝对值和为准则的最优线性组合预测模型为

$$\min F = \sum_{t=1}^{N} |e_t| = \sum_{t=1}^{N} \left| \sum_{i=1}^{m} l_i e_{it} \right|,$$
$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i = 1, \\ l_1, l_2, \cdots, l_m \geqslant 0. \end{cases} \tag{4.3.2}$$

模型 (4.3.2) 可化为如下线性规划问题求得最优解, 令

$$e_t^+ = \begin{cases} e_t, & 当 e_t \geqslant 0 时, \\ 0, & 当 e_t < 0 时, \end{cases} \qquad e_t^- = \begin{cases} 0, & 当 e_t \geqslant 0 时, \\ -e_t, & 当 e_t < 0 时. \end{cases}$$

则有

$$|e_t| = e_t^+ + e_t^-, \quad e_t = e_t^+ - e_t^-.$$

于是模型 (4.3.2) 可化为

$$\min F = \sum_{t=1}^{N} |e_t| = \sum_{t=1}^{N} (e_t^+ + e_t^-),$$
$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i e_{it} - e_t^+ + e_t^- = 0, \\ \sum_{i=1}^{m} l_i = 1, \\ l_1, l_2, \cdots, l_m \geqslant 0. \end{cases} \tag{4.3.3}$$

这是一个线性规划问题, 可以用有现成的线性规划软件, 如 Lingo 来求解.

4.4 以最大误差绝对值达到最小的线性组合预测模型

误差绝对值的大小可以反映预测精度的好坏. 最大误差绝对值达到最小的准则就是预测者根据决策者的要求, 尽量使 N 个时刻中组合预测误差绝对值的最大值尽可能地小作为控制预测精度的标准.

设 e_t 为组合预测在第 t 时刻的预测误差, 则有

$$e_t = x_t - \hat{x}_t = \sum_{i=1}^{m} l_i e_{it}, \quad t = 1, 2, \cdots, N,$$

其中 $e_{it} = (x_t - x_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差, $l_1, l_2, \cdots, l_m$ 分别为 m 种单项预测方法的非负加权系数, 且满足归一性.

设 J_3 表示组合预测误差绝对值的最大值, 则有

$$J_3 = \max_{1 \leqslant t \leqslant N} |e_t|. \tag{4.4.1}$$

因此以最大误差绝对值达到最小的线性组合预测模型可表示为下列最优化问题

$$\min J_3 = \min \max_{1 \leqslant t \leqslant N} |e_t|,$$
$$\text{s.t.} \begin{cases} e_t = \sum_{i=1}^{m} l_i e_{it}, \\ \sum_{i=1}^{m} l_i = 1, \\ l_1, l_2, \cdots, l_m \geqslant 0. \end{cases} \tag{4.4.2}$$

模型 (4.4.2) 可化为如下线性规划问题求得最优解, 令

$$v = \max_{1 \leqslant t \leqslant N} |e_t|.$$

所以 $|e_t| \leqslant v$, $t = 1, 2, \cdots, N$, 则有

$$\min J_3 = v,$$
$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i e_{it} - v \leqslant 0, t = 1, 2, \cdots, N, \\ \sum_{i=1}^{m} l_i e_{it} + v \geqslant 0, t = 1, 2, \cdots, N, \\ \sum_{i=1}^{m} l_i = 1, \\ l_1, l_2, \cdots, l_m \geqslant 0, v \geqslant 0. \end{cases} \tag{4.4.3}$$

这是一个线性规划问题, 有现成的线性规划软件来求解.

4.5 以预测误差全距作为目标函数的组合预测模型

以预测误差平方和达到最小的线性组合预测模型, 要计算 m 阶误差信息矩阵 $\boldsymbol{E}$, 并通过求解一个二次规划模型而获得最优组合预测权系数. 本节考虑以组合预

测误差全距作为目标函数的组合预测模型, 由于该模型为线性规划模型, 因而可以避免一些优化组合模型在确定组合预测权系数中所带来的计算上的困难.

全距反映总体单位中最大样本值与最小样本值之差, 它和方差或标准差一样, 也反映总体某种指标离散程度. 预测者进行组合预测时, 要想提高预测精度, 可采用组合预测误差全距尽量小的预测方法进行组合预测, 剔除一些预测误差全距较大预测方法, 从而使组合预测的精度比较稳定可靠. 下面提出组合预测在不同时刻误差的全距作为控制组合预测精度的一个优化指标, 以极小化全距作为目标函数建立组合预测模型.

假定经过筛选以后, 存在 m 种比较稳定的组合预测方法对某个社会经济现象的某个指标序列进行预测. 设该指标序列为 $\{x_t, t=1,2,\cdots,N\}$, 设 $\hat{x}_t = l_1x_{1t} + l_2x_{2t} + \cdots + l_mx_{mt}$ 为 x_t 的组合预测值, 则 e_t 为在第 t 时刻组合预测的预测误差, 则有

$$e_t = x_t - \hat{x}_t = \sum_{i=1}^{m} l_i e_{it}.$$

显然 e_t 依赖于组合预测的加权系数 $l_1, l_2, \cdots, l_m$, 把 $e_1, e_2, \cdots, e_N$ 按大小顺序排列如下

$$e_{(1)} \leqslant e_{(2)} \leqslant \cdots \leqslant e_{(N)}.$$

即

$$e_{(N)} = \max_{1\leqslant t\leqslant N} e_{(t)}, \quad e_{(1)} = \min_{1\leqslant t\leqslant N} e_{(t)}. \tag{4.5.1}$$

则组合预测的预测误差全距 R 为

$$R = e_{(N)} - e_{(1)}.$$

以组合预测误差全距作为目标函数的组合预测模型表示如下模型 [63]

$$\min R = e_{(N)} - e_{(1)},$$

$$\text{s.t.} \begin{cases} e_{(N)} - \sum_{i=1}^{m} l_i e_{it} \geqslant 0, t = 1,2,\cdots,N, \\ \sum_{i=1}^{m} l_i e_{it} - e_{(1)} \geqslant 0, t = 1,2,\cdots,N, \\ \sum_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1,2,\cdots,m. \end{cases} \tag{4.5.2}$$

我们还可以建立全距作为控制组合预测精度的组合预测模型. 若预测者根据问题的性质, 可确定一个最大允许的组合误差全距 R_0, 在 R_0 的范围内使平均绝对误

差达到最小, 则有如下线性规划模型

$$
\min \sum_{i=1}^{m} l_i \bar{e}_i,
$$

$$
\text{s.t.} \begin{cases} e_{(N)} - e_{(1)} \leqslant R_0, \\ e_{(N)} - \sum\limits_{i=1}^{m} l_i e_{it} \geqslant 0, t = 1, 2, \cdots, N, \\ \sum\limits_{i=1}^{m} l_i e_{it} - e_{(1)} \geqslant 0, t = 1, 2, \cdots, N, \\ \sum\limits_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1, 2, \cdots, m, \end{cases} \tag{4.5.3}
$$

其中 $\bar{e}_i = \dfrac{1}{N} \sum\limits_{t=1}^{N} |e_{it}|$, $\bar{e}_i$ 表示第 i 种单项预测方法的平均绝对误差 $i = 1, 2, \cdots, m$.

模型 (4.5.2)、(4.5.3) 的特点为:

(1) 以预测误差平方和为目标函数的组合预测模型, 要计算 m 阶误差信息矩阵 $\boldsymbol{E}$, 其中

$$
\boldsymbol{E} = \begin{bmatrix} \sum\limits_{t=1}^{N} e_{1t}^2 & \sum\limits_{t=1}^{N} e_{1t} e_{2t} & \cdots & \sum\limits_{t=1}^{N} e_{1t} e_{mt} \\ \sum\limits_{t=1}^{N} e_{2t} e_{1t} & \sum\limits_{t=1}^{N} e_{2t}^2 & \cdots & \sum\limits_{t=1}^{N} e_{2t} e_{mt} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{t=1}^{N} e_{mt} e_{1t} & \sum\limits_{t=1}^{N} e_{mt} e_{2t} & \cdots & \sum\limits_{t=1}^{N} e_{mt}^2 \end{bmatrix}_{m \times m}
$$

当然 $\boldsymbol{E}$ 的运算量比较大. 而模型 (4.5.2)、(4.5.3) 不含 m 阶误差信息矩阵 $\boldsymbol{E}$, 模型 (4.5.3) 也只需要计算 $\bar{e}_i$.

(2) 它们为线性规划模型, 只要引进松弛变量和人工变量就可以用大 M 法来求解, 有现成的线性规划软件来求解. 而以预测误差平方和为目标函数的组合预测模型为非线性规划模型, 其运算量要大于线性规划模型.

(3) 事实上, 预测误差 $\{e_{it}, t = 1, 2, \cdots, N, i = 1, 2, \cdots, m\}$ 可视为随机变量序列, 这是由于指标序列的观测值具有随机性. 当预测误差序列 $\{e_{it}, t = 1, 2, \cdots, N, i = 1, 2, \cdots, m\}$ 服从多元正态分布时, 组合预测模型的预测误差全距和标准差的数学期望之比为常数, 这个常数与样本量有关[64]. 因此, 以预测误差平方和为目标函数的组合预测模型和以组合预测误差全距作为目标函数的组合预测模型的解近似相等.

从统计学的角度可以考虑利用全距来衡量组合预测误差的变动程度, 建立固定权系数的组合预测模型. 在大样本情况下, 由于误差全距与标准差之比为常数, 所以

此类组合预测模型和误差平方和为目标函数组合预测模型具有同等的预测有效性,但计算较简单、实用. 所以以组合预测误差全距作为目标函数的组合预测模型具有一定的推广应用价值.

4.6 非负可变加权系数的组合预测模型

4.6.1 非负变权组合预测模型建立的必要性

目前研究较多的组合预测模型的加权系数是不变的. 事实上, 这样的权系数假定为常数, 组合预测模型可能不能很好地反映预测方法的有效性, 这将会使组合预测精度降低. 产生权系数变化原因有很多, 主要有三个:

一是我国目前的经济体制改革正处于转轨时期, 即从过去的有计划的商品经济向社会主义市场经济转变, 经济系统的各要素的关系发生显著变化. 然而, 这种定结构的组合预测模型在国民经济结构改变的条件下使用可能带来较大的预测误差.

二是不同的预测方法特点不同, 每种预测方法表现出 "时好时坏性", 反映在权重上表现为 "时大时小"[65].

三是不同的预测者对事物的客观规律的认识有差异, 某种预测方法可能随时间的推移越来越优于其他单项预测方法, 从而导致组合预测权系数的变化.

因此, 建立非负变权组合预测模型具有一定的现实意义, 从而可以进一步提高组合预测精度.

文献 [20] 建立可变权系数的组合预测模型. 设有某种社会经济现象的某个指标的时间序列, 存在 m 种单项预测模型 $f_1, f_2, \cdots, f_m$ 对其进行预测, 它们组成的组合预测模型可以表示为

$$f_{ct} = \sum_{i=1}^{m} g_i(t) f_{it},$$

其中 f_{ct} 为第 t 时刻的组合预测值, f_{it} 为第 i 种单项预测模型在第 t 时刻的预测值, $g_i(t)$ 为第 i 种单项预测模型在第 t 时刻的可变加权系数, $i = 1, 2, \cdots, m$, $t = 1, 2, \cdots, N$, 且假定 $g_i(t)$ 为 t 的连续函数, 因此它可以用适当的 n 次多项式来逼近, 即

$$g_i(t) = \sum_{k=0}^{n} g_{ik} t^k, \quad i = 1, 2, \cdots, m.$$

文献 [20] 中考虑了可变加权系数满足如下约束

$$\sum_{k=0}^{n} g_i(t) = 1, \quad t = 1, 2, \cdots, N.$$

但是文献 [20] 中没有考虑了可变加权系数的非负性. 因为加权系数为负值, 其经济含义不够明确. 于是, 本节提出非负可变加权系数的组合预测模型, 并利用数学规划给出了可变加权系数的计算.

4.6.2 最优的非负可变加权系数的组合预测模型的建立

设对同一预测对象的某个指标序列为 $\{x_t, t=1,2,\cdots,N\}$, 存在 m 种单项预测方法对其进行预测, 设 x_{it} 为第 i 种单项预测模型在第 t 时刻的拟合值, $i=1,2,\cdots,m$, $t=1,2,\cdots,N$. 令

$$\hat{x}_t=\sum_{i=1}^{m} l_{it}x_{it}, \tag{4.6.1}$$

其中 $\hat{x}_t$ 第 t 时刻的组合预测值, l_{it} 为第 i 种单项预测方法在第 t 时刻的加权系数, 且 l_{it} 满足

$$\begin{cases}\sum\limits_{i=1}^{m} l_{it}=1, t=1,2,\cdots,N,\\ l_{it}\geqslant 0, i=1,2,\cdots,m, t=1,2,\cdots,N.\end{cases} \tag{4.6.2}$$

设 e_t 为组合预测在第 t 时刻的预测误差, 则有

$$e_t=x_t-\hat{x}_t=\sum_{i=1}^{m} l_{it}e_{it},$$

其中 $e_{it}=(x_t-x_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差.

设 J 表示非负可变加权系数的组合预测的预测误差平方和, 则有

$$J=\sum_{t=1}^{N} e_t^2=\sum_{t=1}^{N}\sum_{i=1}^{m}\sum_{j=1}^{m} l_{it}l_{jt}e_{it}e_{jt}. \tag{4.6.3}$$

因此, 以预测误差平方和为准则的非负可变加权系数的组合预测模型为下列最优化问题

$$\min J=\sum_{t=1}^{N}\sum_{i=1}^{m}\sum_{j=1}^{m} l_{it}l_{jt}e_{it}e_{jt},$$
$$\text{s.t.}\begin{cases}\sum\limits_{i=1}^{m} l_{it}=1, t=1,2,\cdots,N,\\ l_{it}\geqslant 0, i=1,2,\cdots,m, t=1,2,\cdots,N.\end{cases} \tag{4.6.4}$$

记

$$\boldsymbol{L}_t=(l_{1t},l_{2t},\cdots,l_{mt})^{\mathrm{T}},\quad \boldsymbol{E}_t=(e_{1t},e_{2t},\cdots,e_{mt})^{\mathrm{T}}.$$

则

$$e_t = \boldsymbol{L}_t^{\mathrm{T}} \boldsymbol{E}_t, \quad e_t^2 = \boldsymbol{L}_t^{\mathrm{T}} \boldsymbol{E}_t \boldsymbol{E}_t^{\mathrm{T}} \boldsymbol{L}_t,$$

其中 $\boldsymbol{E}_t \boldsymbol{E}_t^{\mathrm{T}}$ 表示 m 种单项预测模型在第 t 时刻预测误差的协方差矩阵.

再令

$$\boldsymbol{L} = \left(\boldsymbol{L}_1^{\mathrm{T}}, \boldsymbol{L}_2^{\mathrm{T}}, \cdots, \boldsymbol{L}_N^{\mathrm{T}}\right)^{\mathrm{T}},$$

$$\boldsymbol{E} = \begin{bmatrix} \boldsymbol{E}_1 \boldsymbol{E}_1^{\mathrm{T}} & & & \\ & \boldsymbol{E}_2 \boldsymbol{E}_2^{\mathrm{T}} & & \\ & & \ddots & \\ & & & \boldsymbol{E}_N \boldsymbol{E}_N^{\mathrm{T}} \end{bmatrix}.$$

显然, $\boldsymbol{L}$ 表示非负可变加权系数向量, $\boldsymbol{E}$ 为准对角矩阵, 表示 m 种单项预测模型在不同时刻预测误差的协方差矩阵. 所以组合预测的预测误差平方和 J 表示为

$$J = \boldsymbol{L}^{\mathrm{T}} \boldsymbol{E} \boldsymbol{L}.$$

令 $\boldsymbol{R} = (1, 1, \cdots, 1)_{1\times m}, \boldsymbol{0} = (0, 0, \cdots, 0)_{1\times m}$, 即 $\boldsymbol{R}$ 是由元素 1 组成的 m 维行向量, $\boldsymbol{0}$ 是由元素 0 组成的 m 维行向量, 设 $\boldsymbol{R}_t$ 是一个 mN 维行向量, 其分块形式如下

$$\boldsymbol{R}_t = (\boldsymbol{0}, \boldsymbol{0}, \cdots, \underset{\substack{\downarrow \\ \text{位于第}t\text{块}}}{\boldsymbol{R}}, \cdots, \boldsymbol{0})_{1\times mN}.$$

则约束条件可以表示为

$$\begin{cases} \boldsymbol{R}_t^{\mathrm{T}} \boldsymbol{L} = 1, \quad t = 1, 2, \cdots, N, \\ \boldsymbol{L} \geqslant \boldsymbol{0}. \end{cases}$$

因此, 以预测误差平方和为准则的非负可变加权系数的组合预测模型可表示成矩阵形式

$$\begin{aligned} &\min J = \boldsymbol{L}^{\mathrm{T}} \boldsymbol{E} \boldsymbol{L}, \\ &\text{s.t.} \begin{cases} \boldsymbol{R}_t^{\mathrm{T}} \boldsymbol{L} = 1, t = 1, 2, \cdots, N, \\ \boldsymbol{L} \geqslant \boldsymbol{0}. \end{cases} \end{aligned} \tag{4.6.5}$$

这是一个二次规划问题. 由二次规划理论知, 此二次规划实际上为一个凸规划. 因此 Kuhn-Tucker 条件既是凸规划最优点存在的必要条件, 也是凸规划最优点存在

的充分条件. 模型 (4.6.5) 的 Kuhn-Tucker 条件可表示为

$$\begin{cases} 2\boldsymbol{EL}-\sum_{t=1}^{N}\lambda_t\boldsymbol{R}_t^{\mathrm{T}}-\boldsymbol{U}=\boldsymbol{0}, \\ \boldsymbol{R}_t^{\mathrm{T}}\boldsymbol{L}=1, t=1,2,\cdots,N, \\ \boldsymbol{U}^{\mathrm{T}}\boldsymbol{L}=0, \\ \boldsymbol{L}\geqslant\boldsymbol{0},\boldsymbol{U}\geqslant\boldsymbol{0}, \end{cases}$$

其中 $\boldsymbol{U}=(u_{11},u_{21},\cdots,u_{m1},u_{12},u_{22},\cdots,u_{m2},\cdots,u_{1N},u_{2N},\cdots,u_{mN})^{\mathrm{T}}$ 是与非负可变组合预测权系数向量 $\boldsymbol{L}$ 所对应的 Kuhn-Tucker 乘子. u_{it} 和 l_{it} 不能同时为基变量, λ_t 是与第 t 个约束条件 $\boldsymbol{R}_t^{\mathrm{T}}\boldsymbol{L}=1$ 所对应的 Lagrange 乘子.

由于 λ_t 无非负约束, 可令 $\lambda_t=\lambda_{1t}-\lambda_{t2}$, $t=1,2,\cdots,N$, 其中 $\lambda_{1t},\lambda_{2t}\geqslant 0$. 因此可构造如下辅助线性规划模型

$$\min J=\sum_{t=1}^{N}v_t,$$

$$\text{s.t.}\begin{cases} 2\boldsymbol{EL}-\sum_{t=1}^{N}(\lambda_{1t}-\lambda_{2t})\boldsymbol{R}_t^{\mathrm{T}}-\boldsymbol{U}=\boldsymbol{0}, \\ \boldsymbol{R}_t^{\mathrm{T}}\boldsymbol{L}+v_t=1, \quad t=1,2,\cdots,N, \\ \boldsymbol{L}\geqslant\boldsymbol{0},\boldsymbol{U}\geqslant\boldsymbol{0}, \\ \lambda_{1t},\lambda_{2t}\geqslant 0, t=1,2,\cdots,N. \end{cases} \tag{4.6.6}$$

解此线性规划模型即可获得非负可变组合预测权系数向量.

4.6.3 以误差绝对值之和达到最小的非负可变加权系数的组合预测模型

设 Q 表示非负可变加权系数的组合预测的误差绝对值之和, 则有

$$Q=\sum_{t=1}^{N}|e_t|=\sum_{t=1}^{N}\left|x_t-\sum_{i=1}^{m}l_{it}x_{it}\right|=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}l_{it}e_{it}\right|. \tag{4.6.7}$$

可得以预测误差绝对值和为准则的线性组合预测模型

$$\min Q=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}l_{it}e_{it}\right|,$$

$$\text{s.t.}\begin{cases} \sum_{i=1}^{m}l_{it}=1, t=1,2,\cdots,N, \\ l_{it}\geqslant 0, i=1,2,\cdots,m, t=1,2,\cdots,N. \end{cases} \tag{4.6.8}$$

令

$$u_t=(|e_t|+e_t)/2, \quad v_t=(|e_t|-e_t)/2,$$

则有

$$|e_t| = u_t + v_t, \quad e_t = u_t - v_t, \quad u_t \cdot v_t = 0, \quad u_t, v_t \geqslant 0, \quad t = 1, 2, \cdots, N.$$

模型 (4.6.8) 可化为如下线性规划问题求得最优解, 于是有

$$\min Q = \sum_{t=1}^{N} (u_t + v_t),$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_{it} e_{it} - u_t + v_t = 0, t = 1, 2, \cdots, N, \\ \sum_{i=1}^{m} l_{it} = 1, t = 1, 2, \cdots, N, \\ u_t \geqslant 0, v_t \geqslant 0, l_{it} \geqslant 0, i = 1, 2, \cdots, m, t = 1, 2, \cdots, N. \end{cases} \tag{4.6.9}$$

这是一个线性规划问题, 有现成的线性规划软件来求解.

4.6.4 以全距达到最小的非负可变加权系数的组合预测模型

设 $e_t = x_t - \hat{x}_t = \sum_{i=1}^{m} l_{it} e_{it}$ 为组合预测在第 t 时刻的预测误差, 把 $e_1, e_2, \cdots, e_N$ 按大小顺序排列如下: $e_{(1)} \leqslant e_{(2)} \leqslant \cdots \leqslant e_{(N)}$, 其中 $e_{(N)} = \max\limits_{1 \leqslant t \leqslant N} e_{(t)}$, $e_{(1)} = \min\limits_{1 \leqslant t \leqslant N} e_{(t)}$, 则组合预测的预测误差全距 $R = e_{(N)} - e_{(1)}$, 所以以全距达到最小的非负可变加权系数的组合预测模型为

$$\min e_{(N)} - e_{(1)},$$

$$\text{s.t.} \begin{cases} e_{(N)} - \sum_{i=1}^{m} l_{it} e_{it} \geqslant 0, t = 1, 2, \cdots, N, \\ \sum_{i=1}^{m} l_{it} e_{it} - e_{(1)} \geqslant 0, t = 1, 2, \cdots, N, \\ \sum_{i=1}^{m} l_{it} = 1, t = 1, 2, \cdots, N, \\ l_{it} \geqslant 0, i = 1, 2, \cdots, m, t = 1, 2, \cdots, N. \end{cases} \tag{4.6.10}$$

这是一个线性规划问题, 有现成的线性规划软件来求解.

4.7 基于预测误差指数的最优组合预测模型的实例分析

4.7.1 组合预测效果评价的指标体系

预测的精确性就是预测的准确度, 它与预测的误差密切相关. 为了反映组合预测效果的好坏, 本节采用以下几种形式的常用的误差指标度量组合预测的准确度.

(1) 预测误差平方和 (SSE)

$$SSE = \sum_{t=1}^{N} (x_t - \hat{x}_t)^2. \tag{4.7.1}$$

(2) 均方误差 (MSE)

$$MSE = \frac{1}{N}\sqrt{\sum_{t=1}^{N} (x_t - \hat{x}_t)^2}. \tag{4.7.2}$$

(3) 平均绝对误差 (MAE)

$$MAE = \frac{1}{N}\sum_{t=1}^{N} |x_t - \hat{x}_t|. \tag{4.7.3}$$

(4) 平均绝对百分比误差 (MAPE)

$$MAPE = \frac{1}{N}\sum_{t=1}^{N} \left|\frac{x_t - \hat{x}_t}{x_t}\right|. \tag{4.7.4}$$

(5) 均方百分比误差 (MSPE)

$$MSPE = \frac{1}{N}\sqrt{\sum_{t=1}^{N} [(x_t - \hat{x}_t)/x_t]^2}. \tag{4.7.5}$$

4.7.2 实例分析

本节以文献 [59] 预测实例说明本章几种最优组合预测方法的应用. 已知用最小二乘法和三次指数平滑法对农村居民储蓄存款余额进行了预测. 其实际观测值 x_t 和两种不同方法的预测值 x_{1t}, x_{2t} 如表 3.4.1 所示.

运用本章的几种最优组合预测模型对此问题进行组合预测, 它们分别是以预测误差平方和达到最小的线性组合预测模型 (4.2.9), 以绝对误差和达到最小的线性组合预测模型 (4.3.2), 以最大误差绝对值达到最小的线性组合预测模型 (4.4.2), 以组合预测误差全距作为目标函数的组合预测模型 (4.5.2). 利用 Matlab 中最优化工具箱对几种最优组合预测模型进行计算, 得到相应的最优组合加权向量 $\boldsymbol{L}^* = (l_1^*, l_2^*)^{\mathrm{T}}$, 结果如表 4.7.1 所示. 为检验组合预测效果的好坏, 表 4.7.1 还同时给出了上述五种组合预测误差的精度分析.

表 4.7.1 几种最优组合预测模型权系数和五种组合预测误差指标

误差预测指标方法	最优权系数	SSE	MSE	MAE	MAPE	MSPE
单项预测方法(1)		401.56	1.67	4.88	0.1959	0.0731
单项预测方法(2)		245.58	1.30	4.59	0.0998	0.0334
组合预测模型式(4.2.9)	$l_1^*=0.4121$ $l_2^*=0.5879$	94.89	0.81	2.34	0.0791	0.0283
组合预测模型式(4.3.2)	$l_1^*=0.3869$ $l_2^*=0.6131$	95.45	0.81	2.30	0.0760	0.0272
组合预测模型式(4.4.2)	$l_1^*=0.3407$ $l_2^*=0.6593$	99.41	0.83	2.39	0.0756	0.0255
组合预测模型式(4.5.2)	$l_1^*=0.3672$ $l_2^*=0.6328$	96.68	0.82	2.3	0.0754	0.0264

从表 4.7.1 可以看出：对于单项预测方法 (1)、(2) 来说，其五种预测误差指标 SSE、MSE、MAE、MAPE、MSPE 均显著地大于几种最优组合预测方法相应的误差指标，这表明最优组合预测方法是优于单项预测方法的. 另外，对于最优组合预测方法来说，组合预测模型 (4.2.9)、(4.3.2) 组合预测效果好，其次是组合预测模型 (4.4.2)、(4.5.2).

第 5 章　基于预测有效度的最优组合预测的有效性理论

目前, 线性组合预测模型的研究多是以误差的平方和达到最小为准则而建立起来的. 对于这类组合预测方法, 其理论研究比较深入[15,66~69], 应用研究也最为广泛. 然而, 基于预测有效度的组合预测模型目前在文献中研究较少. 本章首先给出预测有效度概念的一般表达形式, 其次针对基于一阶或二阶预测有效度的组合预测模型提出了若干概念, 并对其基本问题进行了研究, 主要回答了如下几个基本问题:

第一个基本问题是有关优性组合预测的存在性, 即基于一阶或二阶预测有效度的非负权重最优组合预测方法的预测有效度是否一定大于各个单项预测方法预测有效度中的最大者?

第二个基本问题是有关冗余预测方法的存在性, 即当参加组合的预测方法增多时, 基于一阶或二阶预测有效度的非负权重最优组合预测方法的预测有效度是否一定增大?

第三个基本问题是冗余预测方法的判定, 即给出基于预测有效度的非负权重最优组合预测冗余方法的判定.

5.1　预测有效度的一般数学表达形式

目前, 普通组合预测模型大多以误差的平方和或误差绝对值之和达到最小的准则建立起来的. 事实上, 这样的准则和假定不能很好地反映预测方法的有效性. 原因是不同指标序列有不同量纲, 所以误差的平方和或误差绝对值之和也不具有相同量纲, 不能直接对比. 即使同类指标序列, 量纲相同, 由于同期指标数值不同, 等量的误差的平方和或误差绝对值之和也不能代表预测方法同等有效.

预测方法的有效性除了用预测误差的平方和衡量以外, 还可以用预测精度的均值及反映其离散程度的均方差描述, 但是在某些特定的情况下, 人们要进一步要考虑预测精度分布的偏度和峰度问题. 为此, 文献 [70] 对预测方法的有效性指标作了初探, 它是以 k 阶预测相对误差的无效度元为基础, 定义了一般形式的预测有效度. 鉴于时间序列的离散性, 本节将给出预测有效度等价的一般离散形式.

设某社会经济现象的指标序列的观察值为 $\{x_t, t=1,2,\cdots,N\}$, 设有 m 个单项预测方法对其进行预测, x_{it} 为第 i 种预测方法第 t 时刻的预测值, $i=1,2,\cdots,m, t=$

$1,2,\cdots,N$. 则有如下几个概念.

定义 5.1.1 令

$$e_{it}=\begin{cases}-1, & 当(x_t-x_{it})/x_t<-1时,\\ (x_t-x_{it})/x_t, & 当-1\leqslant(x_t-x_{it})/x_t\leqslant 1时,\\ 1, & 当(x_t-x_{it})/x_t>1时.\end{cases}$$

称 e_{it} 为第 i 种预测方法第 t 时刻的预测相对误差, $i=1,2,\cdots,m,t=1,2,\cdots,N$. 称矩阵 $\boldsymbol{E}=(e_{it})_{m\times N}$ 为组合预测模型的相对误差矩阵.

显然, $0\leqslant|e_{it}|\leqslant 1$, 矩阵 $\boldsymbol{E}$ 的第 i 行为第 i 种预测方法在各个 t 时刻的预测相对误差序列, $\boldsymbol{E}$ 的第 t 列为各种预测方法在第 t 时刻的预测相对误差序列.

定义 5.1.2 称 $A_{it}=1-|e_{it}|$ 为第 i 种预测方法在第 t 时刻的预测精度, $i=1,2,\cdots,m,t=1,2,\cdots,N$. 显然, $0\leqslant A_{it}\leqslant 1$, 当 $|(x_t-x_{it})/x_t|>1$ 时, 则其对应的预测精度 A_{it}=0. 这表明第 i 种预测方法在第 t 时刻的预测为无效预测.

定义 5.1.2 表明, 由于各种因素的影响, e_{it} 具有随机性, 从而 $\{A_{it},i=1,2,\cdots,m,$ $t=1,2,\cdots,N\}$ 可视为一个随机变量序列.

定义 5.1.3 称 $m_i^k=\sum\limits_{t=1}^{N}Q_tA_{it}^k$ 为第 i 种预测方法 k 阶的预测有效度元, k 为正整数, $i=1,2,\cdots,m$, 其中 $\{Q_t,t=1,2,\cdots,N\}$ 为 m 种预测方法在第 t 时刻的离散概率分布, $\sum\limits_{t=1}^{N}Q_t=1$, $Q_t>0$.

特别地, 若对 m 种单项预测方法的预测精度的离散概率分布先验信息不确知时, 可取 $Q_t=1/N,t=1,2,\cdots,N$. 事实上, 第 i 种预测方法的预测有效度元 m_i^k 为第 i 种预测方法预测精度序列 $\{A_{it},t=1,2,\cdots,N\}$ 的 k 阶原点矩.

定义 5.1.4 m_i^k 为第 i 种预测方法 k 阶的预测有效度元, $i=1,2,\cdots,m$, H 为某一 k 元连续函数, 则称 $H(m_i^1,m_i^2,\cdots,m_i^k)$ 为第 i 种预测方法 k 阶预测有效度.

特别地, 引进一阶和二阶预测有效度的概念.

定义 5.1.5 当 $H(x)=x$ 为一元连续函数时, 则 $H(m_i^1)=m_i^1$ 为第 i 种预测方法一阶预测有效度; 当 $H(x,y)=x\left(1-\sqrt{y-x^2}\right)$ 为二元连续函数时, 则 $H\left(m_i^1,m_i^2\right)=m_i^1\left(1-\sqrt{m_i^2-(m_i^1)^2}\right)$ 为第 i 种预测方法二阶预测有效度.

定义 5.1.5 表明一阶预测有效度就是预测精度序列的数学期望. 二阶预测有效度就是预测精度序列的数学期望乘以 1 与其标准差的差. 这正是文献 [71] 提出的组合预测方法有效度的概念. 可见, 定义 5.1.4 对预测有效度的概念作了推广.

5.2　基于一阶预测有效度的组合预测模型

本节给出了预测方法有效度的概念, 在此基础上建立了基于预测方法有效度准则的改进组合预测模型, 并给出了其线性规划的解法.

5.2.1　预测有效度的概念和分类

国内外对预测方法的有效性都进行过一定的分析. 预测方法的有效性应是平均的、全面的或典型的精确性, 所以预测方法的有效性可以用全面的、平均的精度来表达. 也就是说, 某预测方法在某个时期有较高的预测精度, 该预测方法不一定有高的预测有效度. 只有在所有时期都有高的预测精度时, 该预测方法才能有高的预测有效度. 而平均的、全面的精度可用预测精度的均值及反映其离散程度的均方差描述, 本节暂不考虑均方差的影响. 显然, 在预测区间内预测精度的均值越大, 预测方法的有效度越高.

设某种社会经济现象的某个指标的时间序列为 $\{x_t, t=1,2,\cdots,N,N+1,\cdots,N+T\}$, 其中样本区间为 $[1,N]$, 预测区间为 $[N+1,N+T]$. 设有 m 种方法对其进行预测. 设 x_{it} 表示第 i 种预测方法第 t 期拟合值, $t=1,2,\cdots,N, i=1,2,\cdots,m$, 第 i 种预测方法第 t 期拟合相对误差 e_{it} 见定义 5.1.1, 从统计规律看第 i 种预测方法第 t 期拟合相对误差 e_{it} 具有随机变量的性质. 令

$$A_{it}=\begin{cases}1-|e_{it}|, & 当 0\leqslant|e_{it}|\leqslant 1,\\ 0, & 当 |e_{it}|\geqslant 1,\end{cases}$$

称 A_{it} 为第 i 种预测方法在第 t 时刻的预测精度, $i=1,2,\cdots,m, t=1,2,\cdots,N$. 显然预测精度 A_{it} 具有随机变量的性质.

定义 5.2.1[72]　称 $m_{ic}=\sum_{t=1}^{N}Q_{it}A_{it}$ 为样本区间为 $[1,N]$ 上第 i 种预测方法拟合有效度, 其中 Q_{it} 表示第 i 种预测方法在样本区间上第 t 时刻预测精度 A_{it} 的离散概率分布 (或称权重系数), 且满足

$$\sum_{t=1}^{N}Q_{it}=1,\quad Q_{it}>0,\quad t=1,2,\cdots,N.$$

定义 5.2.2　称 $m_{if}=\sum_{t=N+1}^{N+T}Q_{it}A_{it}$ 为预测区间 $[N+1,N+T]$ 上第 i 种预测方法预测有效度, 其中 Q_{it} 表示第 i 种预测方法在预测区间第 t 时刻预测精度 A_{it} 的离散概率分布, 且满足

$$\sum_{t=N+1}^{N+T}Q_{it}=1,\quad Q_{it}>0,\quad t=N+1,N+2,\cdots,N+T.$$

定义 5.2.3 称 $m_i = \alpha m_{ic} + (1-\alpha)m_{if}$ 为整个区间 $[1, N+T]$ 上第 i 种预测方法综合预测有效度, $0 \leqslant \alpha \leqslant 1$, α 越大表示越重视拟合有效度, α 越小表示越重视预测有效度.

在上述三个定义中, m_{ic} 反映了预测模型对样本数据的拟合情况, m_{if} 反映了预测模型的预测结果的有效性, m_i 既能反映预测模型对样本数据的拟合情况, 又能反映预测模型的预测情况. 因而是反映预测方法好坏的综合性指标. 下面分别给出基于不同预测有效度的组合预测模型.

5.2.2 基于预测有效度准则的组合预测模型[72]

1. 在样本区间上组合预测模型权系数的确定

令 $\hat{x}_t = \sum_{i=1}^{m} l_i x_{it}$, 称 $\hat{x}_t$ 为时间序列第 t 时刻组合预测值, $t = 1, 2, \cdots, N$, 其中 $l_1, l_2, \cdots, l_m$ 分别为 m 种单项预测方法在样本区间上的加权系数, $i = 1, 2, \cdots, m$, 加权系数应满足 $l_1 + l_2 + \cdots + l_m = 1, l_i \geqslant 0, i = 1, 2, \cdots, m$. 此时组合预测模型第 t 时刻拟合精度为

$$A_t = 1 - |(x_t - \hat{x}_t)/x_t| = 1 - \left|\sum_{i=1}^{m} l_i e_{it}\right|, \quad t = 1, 2, \cdots, N. \tag{5.2.1}$$

由定义 5.2.1 知, 组合预测模型第 t 时刻拟合有效度为 $m_1 = \sum_{t=1}^{N} Q_t A_t$, 其中 Q_t 表示 A_t 的权系数, 且 $\sum_{t=1}^{N} Q_t = 1$, $Q_t > 0$. 显然, m_1 越大表示组合预测方法越有效. 因此, 以组合预测模型拟合有效度为准则的组合预测模型可表示成如下最优化模型

$$\max m_1 = \sum_{t=1}^{N} Q_t A_t,$$

$$\text{s.t.} \begin{cases} A_t = 1 - \left|\sum_{i=1}^{m} l_i e_{it}\right|, t = 1, 2, \cdots, N, \\ \sum_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{5.2.2}$$

令

$$\varepsilon_t = \sum_{i=1}^{m} l_i e_{it}, \varepsilon_t^+ = \begin{cases} \varepsilon_t, & \text{当}\varepsilon_t \geqslant 0, \\ 0, & \text{当}\varepsilon_t \leqslant 0, \end{cases} \quad \varepsilon_t^- = \begin{cases} -\varepsilon_t, & \text{当}\varepsilon_t \leqslant 0, \\ 0, & \text{当}\varepsilon_t \geqslant 0. \end{cases}$$

则有

$$|\varepsilon_t| = \varepsilon_t^+ + \varepsilon_t^-, \quad \varepsilon_t = \varepsilon_t^+ - \varepsilon_t^-, \quad \varepsilon_t^+ \varepsilon_t^- = 0.$$

注意到 $\sum\limits_{t=1}^{N} Q_t = 1$, 则式 (5.2.2) 等价如下模型

$$
\min z_1 = \sum_{t=1}^{N} Q_t(\varepsilon_t^+ + \varepsilon_t^-),
$$
$$
\text{s.t.} \begin{cases} \sum\limits_{i=1}^{m} l_i e_{it} - \varepsilon_t^+ + \varepsilon_t^- = 0, t = 1, 2, \cdots, N, \\ \sum\limits_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{5.2.3}
$$

模型 (5.2.3) 是一个线性规划问题, 可用大 M 法求解组合预测模型最优加权系数, 有现成的线性规划软件来求最优解 l_i^*, $i = 1, 2, \cdots, m$.

2. 在预测区间上组合预测模型权系数的确定

由定义 5.2.2 知, 组合预测模型在预测区间上预测有效度为

$$
m_2 = \sum_{t=N+1}^{N+T} Q_t A_t,
$$

其中 Q_t 表示 A_t 的已知离散概率分布, 且 $\sum\limits_{t=N+1}^{N+T} Q_t = 1$, $Q_t > 0$.

以预测有效度 m_2 作为目标函数的组合预测模型可表示成如下最优化模型

$$
\max m_2 = \sum_{t=N+1}^{N+T} Q_t A_t,
$$
$$
\text{s.t.} \begin{cases} A_t = 1 - \left| \sum\limits_{i=1}^{m} l_i e_{it} \right|, t = N+1, \cdots, N+T, \\ \sum\limits_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{5.2.4}
$$

令

$$
\varepsilon_t = \sum_{i=1}^{m} l_i e_{it}, \quad \varepsilon_t^+ = \begin{cases} \varepsilon_t, & \text{当}\varepsilon_t \geqslant 0, \\ 0, & \text{当}\varepsilon_t \leqslant 0, \end{cases} \quad \varepsilon_t^- = \begin{cases} -\varepsilon_t, & \text{当}\varepsilon_t \leqslant 0, \\ 0, & \text{当}\varepsilon_t \geqslant 0, \end{cases}
$$
$$
t = N+1, \cdots, N+T.
$$

则有

$$
|\varepsilon_t| = \varepsilon_t^+ + \varepsilon_t^-, \quad \varepsilon_t = \varepsilon_t^+ - \varepsilon_t^-, \quad \varepsilon_t^+ \varepsilon_t^- = 0.
$$

注意到 $\sum_{t=N+1}^{N+T} Q_t = 1$, 则式 (5.2.4) 等价如下模型

$$\min z_2 = \sum_{t=N+1}^{N+T} Q_t(\varepsilon_t^+ + \varepsilon_t^-),$$
$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i e_{it} - \varepsilon_t^+ + \varepsilon_t^- = 0, \quad t = N+1, \cdots, N+T, \\ \sum_{i=1}^{m} l_i = 1, \quad l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{5.2.5}$$

式 (5.2.5) 是一个线性规划模型问题, 可用现成的线性规划软件求解组合预测模型最优加权系数.

3. *在总区间上组合预测模型权系数的确定*

由定义 5.2.3 知, 在总区间上组合预测模型的综合预测有效度为

$$m_3 = \alpha m_1 + (1-\alpha) m_2 = \alpha \sum_{t=1}^{N} Q_t A_t + (1-\alpha) \sum_{t=N+1}^{N+T} Q_t A_t,$$

其中 Q_t 表示 A_t 的已知离散概率分布, 且 $\sum_{t=1}^{N} Q_t = 1$, $\sum_{t=N+1}^{N+T} Q_t = 1, Q_t > 0$.

以综合预测有效度 m_3 作为目标函数的组合预测模型可表示成如下模型

$$\max m_3 = \alpha \sum_{t=1}^{N} Q_t A_t + (1-\alpha) \sum_{t=N+1}^{N+T} Q_t A_t,$$
$$\text{s.t.} \begin{cases} A_t = 1 - \left| \sum_{i=1}^{m} l_i e_{it} \right|, t = 1, 2, \cdots, N, N+1, \cdots, N+T, \\ \sum_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{5.2.6}$$

令

$$\varepsilon_t = \sum_{i=1}^{m} l_i e_{it}, \quad \varepsilon_t^+ = \begin{cases} \varepsilon_t, & \text{当}\varepsilon_t \geqslant 0, \\ 0, & \text{当}\varepsilon_t \leqslant 0, \end{cases}$$

$$\varepsilon_t^- = \begin{cases} -\varepsilon_t, & \text{当}\varepsilon_t \leqslant 0, \\ 0, & \text{当}\varepsilon_t \geqslant 0, \end{cases} \quad t = 1, 2, \cdots, N, N+1, \cdots, N+T.$$

则式 (5.2.6) 等价如下模型

$$\min z_3 = \alpha \sum_{t=1}^{N} Q_t(\varepsilon_t^+ + \varepsilon_t^-) + (1-\alpha) \sum_{t=N+1}^{N+T} Q_t(\varepsilon_t^+ + \varepsilon_t^-),$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i e_{it} - \varepsilon_t^+ + \varepsilon_t^- = 0, t = 1, 2, \cdots, N, N+1, \cdots, N+T, \\ \sum_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{5.2.7}$$

基于预测方法有效度准则的组合预测模型的特点：

(1) 改进组合预测模型的目标函数是线性的, 可用线性规划方法来求解. 而以误差的平方和达到最小的为准则建立起来的组合预测模型, 只能用二次规划方法来求解. 显然, 二次规划求解运算量较大, 从而具有运算简便的特点.

(2) 从改进组合预测模型的等价模型的目标函数上看, 其目标函数为组合预测相对误差的绝对值之和的加权平均, 因此, 基于预测方法有效度准则的改进组合预测模型消除了目标函数的量纲, 从而使不同时间序列组合预测模型有效性具有可比性.

本节分别在样本区间、预测区间及总区间上以不同的有效度指标为准则, 建立改进组合预测模型, 从而区分拟合、估计、预测的概念.

5.3　基于一阶预测有效度的非劣性组合预测和优性组合预测存在的条件

组合预测模型根据建立的某个准则的优劣程度, 可以分为非劣性组合预测、优性组合预测. 文献 [18] 针对无非负约束的以误差平方和达到最小的最优组合预测模型, 提出了非劣性组合预测、优性组合预测的概念. 文献 [70, 71] 提出了组合预测方法有效度的概念, 该指标以预测精度来反映预测方法的有效性, 具有一定的合理性. 为此, 文献 [72] 提出了基于预测有效度的组合预测模型, 并给出了其线性规划的解法. 由于在该模型中不存在组合预测绝对误差信息矩阵, 所以文献 [18] 提出的非劣性组合预测、优性组合预测的概念不再适用, 有必要修改这三个概念. 本节在文献 [72] 的基础上提出了优性组合预测、预测方法优超等概念.

5.3.1　基于一阶预测有效度的组合预测模型的几个概念[73]

设 $\hat{x}_t = l_1 x_{1t} + l_2 x_{2t} + \cdots + l_m x_{mt}$ 为 x_t 的组合预测值, $l_1, l_2, \cdots, l_m$ 为各种预测方法的加权系数, 且满足 $\sum_{i=1}^{m} l_i = 1$, $l_i \geqslant 0$, $i = 1, 2, \cdots, m$.

设 A_{ct} 为第 t 时刻的组合预测精度, 则有

$$A_{ct}=1-|(x_t-\hat{x}_t)/x_t|=1-\left|\sum_{i=1}^{m}l_i\left(x_t-x_{it}\right)/x_t\right|=1-\left|\sum_{i=1}^{m}l_ie_{it}\right|,t=1,2,\cdots,N,$$

其中, $e_{it}=(x_t-x_{it})/x_t$ 表示相对误差.

根据定义5.1.5,一阶组合预测有效度 m_c为其预测精度序列 $\{A_{ct},t=1,2,\cdots,N\}$ 的数学期望. 即

$$m_c=\sum_{t=1}^{N}Q_tA_{ct},$$

其中 $\{Q_t,t=1,2,\cdots,N\}$ 为已知的离散概率分布, 显然, $0\leqslant m_c\leqslant 1$.

m_c 越大表示组合预测方法越有效, 因此, 以一阶组合预测有效度为准则的组合预测模型可表示成如下模型 [72]

$$\max m_c=\sum_{t=1}^{N}Q_tA_{ct},$$
$$\text{s.t.}\begin{cases}A_{ct}=1-\left|\displaystyle\sum_{i=1}^{m}l_ie_{it}\right|,\\ \displaystyle\sum_{i=1}^{m}l_i=1,l_i\geqslant 0,i=1,2,\cdots,m.\end{cases}\tag{5.3.1}$$

记 $m_{\min}=\min\{m_i,i=1,2,\cdots,m\}$, $m_{\max}=\max\{m_i,i=1,2,\cdots,m\}$. 即 $m_{\min}$ 表示 m 种预测方法最小预测有效度, $m_{\max}$ 表示 m 种预测方法最大预测有效度, 则有如下定义.

定义 5.3.1 若 $m_{\min}\leqslant m_c\leqslant m_{\max}$, 则称组合预测模型 (5.3.1) 为非劣性组合预测. 若 $m_c>m_{\max}$, 则称组合预测模型 (5.3.1) 为优性组合预测.

定义 5.3.1 表明, 只有组合预测有效度大于各单项预测有效度中最大者, 则该组合预测模型才为优性组合预测.

定义 5.3.2 若组合预测模型的相对误差矩阵 $\boldsymbol{E}$ 的第 i 行和第 k 行元素满足不等式 $|e_{it}|\leqslant|e_{kt}|$, $t=1,2,\cdots,N$, 且至少对某个时刻 t_0 有严格的不等号成立, $t_0\in\{1,2,\cdots,N\}$, 则称第 i 种单项预测方法优超第 k 种单项预测方法.

定义 5.3.2 表明, 第 i 种单项预测方法在各个时刻预测相对误差的绝对值不大于第 k 种单项预测方法, 而且至少对某个时刻 t_0 处预测相对误差的绝对值严格小于第 k 种单项预测方法, 直观上可以认为第 i 种单项预测方法要 "好于" 第 k 种单项预测方法.

定义 5.3.3 若某种单项预测方法增加到组合预测模型中不能使一阶组合预测有效度得到增加, 则称该单项预测方法为冗余预测方法. 即该种单项预测方法在组合预测模型最优权系数中为零, 表明该种单项预测方法只提供冗余信息.

定义 5.3.4 在一个组合预测模型中, 设共有 m 种单项预测方法参与组合预测, 若最优解中出现冗余预测方法的个数为 m', 则称比例系数 $k=\dfrac{m'}{m}$ 为组合预测模型的冗余度.

显然 $0\leqslant k\leqslant\dfrac{m-1}{m}$. $k=0$ 表示 m 种单项预测方法在一个组合预测模型中均提供有效信息. $k=\dfrac{m-1}{m}$ 表示在一个组合预测模型中只有一个单项预测方法提供有效信息, 而其他 $(m-1)$ 个单项预测方法均为冗余预测方法. 所以冗余度 k 越小表示组合预测模型选择的单项预测方法越有效.

5.3.2 基于一阶预测有效度的非劣性组合预测和优性组合预测存在的条件[73]

1. 简单平均组合预测方法是优性组合预测的条件

简单平均组合预测方法也称为等权平均法, 它是研究较多的正权系数组合预测方法之一. 一般地说, 简单平均组合预测方法不是优性组合预测, 更不用说是最优性组合预测. 但是, 下面的结论指出简单平均组合预测方法至少是非劣性组合预测, 并且在一些条件下它可能成为优性组合预测和最优性组合预测.

先介绍一个引理.

引理 5.3.1 组合预测有效度 m_c 满足如下不等式

$$m_c\geqslant\sum_{i=1}^{m}l_im_i. \tag{5.3.2}$$

证明 因为$\left|\sum\limits_{i=1}^{m}l_ie_{it}\right|\leqslant\sum\limits_{i=1}^{m}l_i\left|e_{it}\right|$, 且注意到$\sum\limits_{i=1}^{m}l_i=1,l_i\geqslant0,\sum\limits_{t=1}^{N}Q_t=1,Q_t>0.$

$$\begin{aligned}m_c&=\sum_{t=1}^{N}Q_tA_{ct}=\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m}l_ie_{it}\right|\right)\\&\geqslant\sum_{t=1}^{N}Q_t\left(1-\sum_{i=1}^{m}l_i\left|e_{it}\right|\right)=\sum_{t=1}^{N}Q_t\sum_{i=1}^{m}l_i\left(1-\left|e_{it}\right|\right)\\&=\sum_{i=1}^{m}l_i\sum_{t=1}^{N}Q_t\left(1-\left|e_{it}\right|\right)=\sum_{i=1}^{m}l_i\sum_{t=1}^{N}Q_tA_{it}=\sum_{i=1}^{m}l_im_i.\end{aligned}$$

即 $m_c\geqslant\sum\limits_{i=1}^{m}l_im_i$. 证毕.

引理 5.3.1 表明, 组合预测有效度大于或等于各单项预测模型的预测有效度的加权平均值.

定理 5.3.1 组合预测模型 (5.3.1) 的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 设 $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为组合预测模型 (5.3.1) 的任一个可行解, 则有

$$\sum_{i=1}^{m} l_i=1,\quad l_i \geqslant 0,\quad i=1,2,\cdots,m.$$

由式 (5.3.2) 知

$$m_c \geqslant \sum_{i=1}^{m} l_i m_i \geqslant \sum_{i=1}^{m} l_i m_{\min}=m_{\min}.$$

由定义 5.3.1 得结论成立. 证毕.

定理 5.3.1 表明, 从组合预测有效度角度而言, 组合预测模型 (5.3.1) 的任一个归一化非负权系数所对应的组合预测均不会比 "最差" 的单项预测方法还要差.

推论 5.3.1 简单平均组合预测方法至少是非劣性组合预测.

证明 因为在简单平均组合预测方法中, 即令组合预测权系数 $l_1=l_2=\cdots=l_m=\dfrac{1}{m}$, 显然它们是组合预测模型 (5.3.1) 的一个可行解, 由定理 5.3.1 知结论成立. 证毕.

定理 5.3.2 在组合预测模型 (5.3.1) 的相对误差矩阵 $\boldsymbol{E}=(e_{it})_{m\times N}$ 中, 若其任意的第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt},t=1,2,3,\cdots,N$ 的符号不全相同, 则组合预测模型 (5.3.1) 的最优解对应的组合预测一定为优性组合预测方法.

证明 因为 $\boldsymbol{E}$ 中任意的第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号不全相同, 所以

$$\left|\sum_{i=1}^{m} l_i e_{it}\right|<\sum_{i=1}^{m} l_i\left|e_{it}\right|,\quad t=1,2,3,\cdots,N.$$

从而

$$\begin{aligned} m_c &= \sum_{t=1}^{N} Q_t A_{ct}=\sum_{t=1}^{N} Q_t\left(1-\left|\sum_{i=1}^{m} l_i e_{it}\right|\right)>\sum_{t=1}^{N} Q_t\left(1-\sum_{i=1}^{m} l_i\left|e_{it}\right|\right) \\ &= \sum_{i=1}^{m} l_i \sum_{t=1}^{N} Q_t A_{it}=\sum_{i=1}^{m} l_i m_i, \end{aligned}$$

即 $m_c>\sum_{i=1}^{m} l_i m_i$. 设 $\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型 (5.3.1) 的最优解, m_c^* 为 $\boldsymbol{L}^*$ 对应的一阶组合预测有效度, 则有

$$m_c^* \geqslant m_c>\sum_{i=1}^{m} l_i m_i.$$

因为 $(1,0,\cdots,0)^{\mathrm{T}},(0,1,\cdots,0)^{\mathrm{T}},\cdots,(0,0,\cdots,1)^{\mathrm{T}}$. 这 m 个 m 维列向量分别是组合预测模型 (5.3.1) 的可行解, 所以 $m_c^*>m_i,i=1,2,\cdots,m$. 即

$$m_c^*>\max\{m_i,i=1,2,\cdots,m\}=m_{\max}.$$

再由定义 5.3.1 得结论成立. 证毕.

定理 5.3.2 表明, 在组合预测模型 (5.3.1) 的相对误差信息矩阵 $\boldsymbol{E}$ 中, 只要某一个时刻 m 种单项预测方法的相对误差有正有负, 则组合预测模型的相对误差减小, 从而组合预测有效度就会提高. 组合预测模型 (5.3.1) 的最优解对应的组合预测一定比 "最好" 的单项预测方法还要 "好".

定理 5.3.3　若组合预测模型 (5.3.1) 的相对误差矩阵 $\boldsymbol{E}=(e_{it})_{m\times N}$ 的秩为 1, $\boldsymbol{E}$ 的第 1 列存在非零元素, 且满足 $\sum\limits_{i=1}^{m} e_{i1}=0$, 则简单平均组合预测方法为最优性组合预测方法. 即 $l_1=l_2=\cdots=l_m=1/m$ 为组合预测模型 (5.3.1) 的最优解.

证明　对 $\boldsymbol{E}$ 按列进行分块, $\boldsymbol{e}_t=(e_{1t},e_{2t},\cdots,e_{mt})^{\mathrm{T}}$ 为 $\boldsymbol{E}$ 的第 t 列, $t=1,2,\cdots,N$. 因为 $\boldsymbol{E}=(e_{it})_{m\times N}$ 的秩为 1, 所以存在比例常数 k_t, 使得 $\boldsymbol{e}_t=k_t\boldsymbol{e}_1$, $t=2,3,\cdots,N$. 又因为 $\sum\limits_{i=1}^{m} e_{i1}=0$, 所以 $\sum\limits_{i=1}^{m} e_{it}=k_t\sum\limits_{i=1}^{m} e_{i1}=0, t=2,3,\cdots,N$. 当简单组合预测方法的权系数 $l_1=l_2=\cdots=l_m=\dfrac{1}{m}$ 时, 则有

$$m_c=\sum_{t=1}^{N}Q_tA_{ct}=\sum_{t=1}^{N}Q_t\left(1-\left|\frac{1}{m}\sum_{i=1}^{m}e_{it}\right|\right)=\sum_{t=1}^{N}Q_t[1-0]=1.$$

即 m_c 达到其上界. 所以简单平均组合预测方法为最优性组合预测方法, 当然它也是优性组合预测方法. 证毕.

2.　基于预测有效度的优性组合预测存在的充分条件

下面给出基于预测有效度的组合预测模型 (5.3.1) 的优性组合预测存在的一个充分条件.

定理 5.3.4　设 $m_1=\max\{m_i, i=1,2,\cdots,m\}$, 即 m_1 表示 m 种单项预测方法中的最大一阶预测有效度, $b_t\in(-|e_{1t}|,|e_{1t}|)$, $t=1,2,\cdots,N$, 若如下 $(N+1)\times m$ 的线性方程组

$$\begin{cases}e_{11}l_1+e_{21}l_2+\cdots+e_{m1}l_m=b_1,\\ e_{12}l_1+e_{22}l_2+\cdots+e_{m2}l_m=b_2,\\ \qquad\cdots\cdots\cdots\cdots\\ e_{1N}l_1+e_{2N}l_2+\cdots+e_{mN}l_m=b_N,\\ l_1+l_2+\cdots+l_m=1\end{cases}\tag{5.3.3}$$

存在非负解, 则组合预测模型 (5.3.1) 的一定存在优性组合预测.

证明　设如上的 $(N+1)\times m$ 的线性方程组存在的非负解为 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$, 则有

$$\sum_{i=1}^{m}e_{it}l_i^*=b_t,\quad t=1,2,\cdots,N,\quad \sum_{i=1}^{m}l_i^*=1,\quad l_i^*\geqslant 0.$$

因为 $b_t \in (-|e_{1t}|, |e_{1t}|)$, 所以 $\left|\sum_{i=1}^{m} e_{it} l_i^*\right| < |e_{1t}|, t = 1, 2, \cdots, N$. 从而 $\sum_{t=1}^{N} Q_t \sum_{i=1}^{m} |l_i^* e_{it}| < \sum_{t=1}^{N} Q_t |e_{1t}|$. 且注意到 $\sum_{t=1}^{N} Q_t = 1$, $Q_t > 0$, $t = 1, 2, \cdots, N$, 则有

$$\begin{aligned} m_c &= \sum_{t=1}^{N} Q_t A_{ct} = \sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i^* e_{it}\right|\right) > \sum_{t=1}^{N} Q_t (1 - |e_{1t}|) = \sum_{t=1}^{N} Q_t A_{1t} \\ &= m_1 = \max\{m_i, i = 1, 2, \cdots, m\}. \end{aligned}$$

即 $m_c > \max\{m_i, i = 1, 2, \cdots, m\}$, 由定义 5.3.1 知结论成立. 证毕.

推论 5.3.2 设 $m_1 = \max\{m_i, i = 1, 2, \cdots, m\}$, 即 m_1 表示 m 种单项预测方法中的最大一阶预测有效度, 若组合预测模型的相对误差矩阵中任一列 m 个元素 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的算术平均数 $\bar{e}_t = \frac{1}{m}\sum_{i=1}^{m} e_{it}$ 满足 $\bar{e}_t \in (-|e_{1t}|, |e_{1t}|), t = 1, 2, \cdots, N$, 则简单平均组合预测方法是优性组合预测方法.

证明 令 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}} = \left(\frac{1}{m}, \frac{1}{m}, \cdots, \frac{1}{m}\right)^{\mathrm{T}}$, 由条件 $\bar{e}_t = \frac{1}{m}\sum_{i=1}^{m} e_{it} \in (-|e_{1t}|, |e_{1t}|)$ 知, 它是满足于线性方程组 (5.3.3) 的非负解, 从而由定理 5.3.4 得推论成立. 证毕.

5.3.3 实例分析

设两种单项预测方法分别对某个经济指标的五年数据进行预测, 则得相对误差矩阵为

$$\boldsymbol{E} = (e_{it})_{2\times 5} = \begin{bmatrix} 5\% & 6\% & 4\% & 7\% & 8\% \\ -6\% & -7\% & -8\% & -9\% & -10\% \end{bmatrix}.$$

若这两个单项预测方法在不同时刻预测精度的加权系数先验信息不确知时, 可以认为预测精度的分布为均匀分布, 即 $Q_t = 1/5, t = 1, 2, \cdots, 5$, 则由定义 5.1.5 知, 第一种和第二种单项预测方法的一阶预测有效度为 $m_1 = \sum_{t=1}^{5} Q_t A_{1t} = 0.94, m_2 = \sum_{t=1}^{5} Q_t A_{2t} = 0.92$. 所以有 $m_1 > m_2$. 设 l_1, l_2 分别为组合预测中两个单项预测方法的加权系数, 且满足

$$l_1 + l_2 = 1, \quad l_i \geqslant 0, \quad i = 1, 2.$$

考察如下 6×2 的线性方程组

$$\begin{cases} 5\%l_1-6\%l_2=b_1, \\ 6\%l_1-7\%l_2=b_2, \\ 4\%l_1-8\%l_2=b_3, \\ 7\%l_1-9\%l_2=b_4, \\ 8\%l_1-10\%l_2=b_5, \\ l_1+l_2=1, \end{cases} \tag{5.3.4}$$

其中 $b_1\in(-5\%,5\%),b_2\in(-6\%,6\%),b_3\in(-4\%,4\%),b_4\in(-7\%,7\%),b_5\in(-8\%,8\%)$. 对线性方程组 (5.3.4), 令 $l_1=0.5$, $l_2=0.5$, 则它们是满足该线性方程组的非负解, 由定理 5.3.4 知组合预测模型 (5.3.1) 一定存在优性组合预测.

事实上, 通过简单的计算, 当 $l_1=0.5, l_2=0.5$ 时, 其对应的组合预测模型 (5.3.1) 预测有效度 $m_{\rm c}=0.984$, 所以 $m_{\rm c}>m_{\max}=\max\{m_1,m_2\}$. 这个例子验证了定理 5.3.4 的结论.

本节关于优性组合预测方法的研究, 实际上是解决了基于一阶预测有效度的组合预测的第一个基本问题. 即非负权重最优组合预测方法的预测有效度不一定大于各个单项预测方法预测有效度中的最大者, 可能出现等于的情况. 也就是说优性组合预测存在是有条件的.

5.4 基于一阶预测有效度组合预测方法冗余信息的判定

若某种单项预测方法增加到组合预测模型中, 它在组合预测模型最优权系数中为零, 这表明该种单项预测方法只提供冗余信息. 下面的定理 5.4.1 说明了冗余信息产生的缘由.

对于组合预测模型 (5.3.1) 有如下结论成立.

定理 5.4.1 组合预测模型 (5.3.1) 的最优目标函数值是 m 的单调不减函数, 即

$$m_c^*(l_1,l_2,\cdots,l_m)\leqslant m_c^*(l_1,l_2,\cdots,l_{m+1}),$$

其中 m 为参与组合预测的各单项预测方法总个数, $m_c^*(l_1,l_2,\cdots,l_m)$ 表示 m 个单项预测方法参与的组合预测模型 (5.3.1) 对应的最大一阶组合预测有效度, $m_c^*(l_1,l_2,\cdots,l_{m+1})$ 表示再增加一个单项预测方法参与的组合预测模型 (5.3.1) 对应的最大一阶组合预测有效度.

证明 设 $\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\rm T}$ 为 m 个单项预测方法参与的组合预测模型

(5.3.1) 的最优解, 则有

$$m_c^*(l_1,l_2,\cdots,l_m)=\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m}l_i^*e_{it}\right|\right),$$

其中 $\sum_{i=1}^{m}l_i^*=1, l_i^*\geqslant 0, i=1,2,\cdots,m.$

同理, 设 $\bar{\boldsymbol{L}}=(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})^{\mathrm{T}}$ 为再增加一个单项预测方法, 共 $m+1$ 个单项预测方法参与的组合预测模型 (5.3.1) 的最优解, 则有

$$m_c^*(l_1,l_2,\cdots,l_m,l_{m+1})=\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m+1}\bar{l}_ie_{it}\right|\right),$$

其中 $\sum_{i=1}^{m+1}\bar{l}_i=1, \bar{l}_i\geqslant 0,\ i=1,2,\cdots,m+1.$

令 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*,0)^{\mathrm{T}}$, 显然, $\boldsymbol{L}$ 为 $m+1$ 个单项预测方法参与的组合预测模型 (5.3.1) 的可行解, 则有

$$\begin{aligned}m_c^*(l_1,l_2,\cdots,l_m,l_{m+1})&\geqslant m_c(l_1^*,l_2^*,\cdots,l_m^*,0)\\&=\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m}l_i^*e_{it}\right|\right)=m_c^*(l_1,l_2,\cdots,l_m).\end{aligned}$$

从而结论成立. 证毕.

一般认为, 随着参与组合预测的单项预测方法个数 m 的增加, 组合预测模型 (5.3.1) 的最大一阶组合预测有效度一定是 m 的严格单调增加函数, 然而定理 5.4.1 证明了当再增加一个单项预测方法时, 组合预测模型 (5.3.1) 的最大一阶组合预测有效度可能不变. 这表明组合预测模型 (5.3.1) 可能存在冗余预测方法. 下面两个定理为冗余信息提供了判定.

定理 5.4.2 若组合预测模型的相对误差矩阵 $\boldsymbol{E}=(e_{it})_{m\times N}$ 中, 任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t=1,2,\cdots,N$, 设 $m_1=\max\{m_i,i=1,2,\cdots,m\}$, 则组合预测模型 (5.3.1) 的冗余度为 $\dfrac{m-1}{m}$, 即后面 $m-1$ 种单项预测方法均为冗余预测方法.

证明 因为 $\boldsymbol{E}$ 中任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t\in\{1,2,\cdots,N\}$. 所以 $\left|\sum_{i=1}^{m}l_ie_{it}\right|=\sum_{i=1}^{m}l_i\left|e_{it}\right|, t\in\{1,2,\cdots,N\}$. 类似于引理的证明, 并注意到

$$\sum_{i=1}^{m}l_i=1,\quad l_i\geqslant 0,\quad \sum_{t=1}^{N}Q_t=1,\quad Q_t>0,\quad t=1,2,\cdots,N.$$

从而

$$\begin{aligned}m_c&=\sum_{t=1}^{N}Q_tA_{ct}=\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m}l_ie_{it}\right|\right)=\sum_{t=1}^{N}Q_t\left(1-\sum_{i=1}^{m}l_i\left|e_{it}\right|\right)\\&=\sum_{t=1}^{N}Q_t\sum_{i=1}^{m}l_i\left(1-\left|e_{it}\right|\right)=\sum_{i=1}^{m}l_i\sum_{t=1}^{N}Q_t\left(1-\left|e_{it}\right|\right)\\&=\sum_{i=1}^{m}l_i\sum_{t=1}^{N}Q_tA_{it}=\sum_{i=1}^{m}l_im_i.\end{aligned}$$

由于 $m_1=\max\{m_i,i=1,2,\cdots,m\}$, 所以 $m_c=\sum\limits_{i=1}^{m}l_im_i\leqslant m_1$.

即组合预测模型 (5.3.1) 的最优解为 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}=(1,0,\cdots,0)^{\mathrm{T}}$, 所以后面 $m-1$ 种单项预测方法均为冗余预测方法. 证毕.

定理 5.4.2 表明, 若组合预测模型在任意第 t 时刻的相对误差符号一致, 即各单项预测方法提供的信息与第一种单项预测方法相似, 则一阶组合预测有效度实际上等于各单项预测方法预测有效度的线性组合, 从而最优的一阶组合预测有效度为第一种单项预测方法最大的预测有效度. 即后面 $m-1$ 种单项预测方法均为冗余预测方法.

定理 5.4.3 若组合预测模型的相对误差矩阵 $\boldsymbol{E}=(e_{it})_{m\times N}$ 中, 任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t=1,2,\cdots,N$, 且第 i 种单项预测方法优超第 k 种单项预测方法, 则组合预测模型 (5.3.1) 的冗余度至少为 $1/m$. 即至少存在一种单项预测方法为冗余预测方法.

证明 因为 $\boldsymbol{E}$ 中任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t\in\{1,2,\cdots,N\}$, 所以在定理 5.4. 2 中已证明得 $m_c=\sum\limits_{i=1}^{m}l_im_i$.

因为第 i 种单项预测方法优超第 k 种单项预测方法, 所以由定义 5.3.2 知 $|e_{it}|\leqslant|e_{kt}|,t=1,2,\cdots,N$, 且至少对某个 t_0 有严格的不等号成立, $t_0\in\{1,2,\cdots,N\}$. 所以 $A_{it}\geqslant A_{kt}$, $t=1,2,\cdots,N$. 且至少对某个 t_0 有严格的不等号成立, 即 $A_{it_0}>A_{kt_0}$, 则

$$m_i=\sum_{t=1}^{N}Q_tA_{it}>\sum_{t=1}^{N}Q_tA_{kt}=m_k.$$

即 $m_i>m_k$. 假设第 k 种单项预测方法不为冗余预测方法, 设组合预测模型 (5.3.1) 的最优解为

$$\boldsymbol{L}^*=(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}},\quad l_k^*\neq0,\quad \sum_{i=1}^{m}l_i^*=1.$$

$\boldsymbol{L}^*$ 对应的目标函数值为

$$m_c = l_1^* m_1 + \cdots + l_i^* m_i + \cdots + l_k^* m_k + \cdots + l_m^* m_m.$$

构造组合预测模型 (5.3.1) 的另一可行解为 $\hat{\boldsymbol{L}} = (l_1^*, \cdots, l_i^* + l_k^*, \cdots, 0, \cdots, l_m^*)^{\mathrm{T}}$, $\hat{\boldsymbol{L}}$ 对应的目标函数值为

$$\hat{m}_c = l_1^* m_1 + \cdots + (l_i^* + l_k^*) m_i + \cdots + 0 \times m_k + \cdots + l_m^* m_m.$$

由 $m_i > m_k$ 知 $\hat{m}_c > m_c$, 而这与 $\boldsymbol{L}^* = (l_1^*, \cdots, l_i^*, \cdots, l_k^*, \cdots, l_m^*)^{\mathrm{T}}$ 为组合预测模型 (5.3.1) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 此即组合预测模型 (5.3.1) 的冗余度至少为 $1/m$. 证毕.

定理 5.4.3 表明, 若第 i 种单项预测方法优超第 k 种单项预测方法, 则可以把组合预测模型相对误差矩阵 $\boldsymbol{E}$ 的第 k 行删除后, 从剩下的较低阶相对误差矩阵所对应的组合预测模型求最优组合权向量, 从而可以简化计算.

本节关于冗余信息的判定的研究, 实际上是解决了基于预测有效度的组合预测的第二个基本问题和第三个基本问题. 即当参加组合的预测方法增多时, 非负权重最优组合预测方法的预测有效度不一定严格增加, 冗余信息完全可能出现. 同时也给出冗余信息出现的两种情形.

5.5 基于二阶预测有效度的优性组合预测模型

前面已针对基于一阶预测有效度的组合预测模型, 提出了新的劣性组合预测、非劣性组合预测、优性组合预测、预测方法优超和冗余度等概念, 给出了非劣性组合预测的条件、简单平均方法是优性组合预测存在的充分条件和优性组合预测存在的充分条件, 同时给出了组合预测模型中冗余信息出现的两个判定定理. 由于前面只把预测精度的数学期望作为一阶预测有效度, 而没有考虑预测精度的方差对预测方法的有效性的影响. 实际上预测方法的有效性应是平均的、全面的精确性, 而平均的、全面的精度可用精度的均值及反映其离散程度的标准差描述, 即在预测区间内预测精度的均值越大, 预测方法的有效度越高, 而预测精度序列的离散程度越大, 预测方法的有效度越低. 即预测方法的有效性与预测精度的均值成正比, 与其标准差成反比. 所以有必要进一步研究基于二阶预测有效度的组合预测方法, 以丰富组合预测理论.

本节首先建立考虑预测精度标准差的预测有效度的组合预测模型, 然后针对该模型重新定义预测有效度. 在提出非劣性组合预测、优性组合预测、预测方法优超等概念的基础上. 继续探讨了简单平均方法是优性组合预测的条件以及优性组合预测存在的一个充分条件. 并给出了冗余信息出现的判定定理. 最后也讨论了组合预测模型的新的近似计算方法, 给出实例分析.

5.5.1 几个推广的概念[74]

设某社会经济现象的指标序列的观察值为 $\{x_t, t=1,2,\cdots,N\}$, 设有 m 个单项预测方法对其进行预测, x_{it} 为第 i 种预测方法第 t 时刻的预测值, $i=1,2,\cdots,m, t=1,2,\cdots,N$. e_{it} 为第 i 种预测方法第 t 时刻的预测相对误差, $A_{it}=1-|e_{it}|$ 为第 i 种预测方法在第 t 时刻的预测精度, $i=1,2,\cdots,m, t=1,2,\cdots,N$. 则有如下几个概念.

定义 5.5.1 称 $M_i=E(A_i)(1-\sigma(A_i))$ 为第 i 种预测方法的二阶预测有效度, $i=1,2,\cdots,m$. 其中 $E(A_i)=\sum_{t=1}^{N}Q_tA_{it}, \sigma(A_i)=\left[\sum_{t=1}^{N}Q_t\left(A_{it}-E(A_i)\right)^2\right]^{\frac{1}{2}}, \{Q_t, t=1,2,\cdots,N\}$ 为 m 种预测方法在第 t 时刻的离散概率分布, $\sum_{t=1}^{N}Q_t=1, Q_t>0$. 即 $E(A_i)$ 为第 i 种预测方法的预测精度序列的数学期望, $\sigma(A_i)$ 为第 i 种预测方法的预测精度序列的标准差.

显然, 定义 5.5.1 与定义 5.1.5 的二阶预测有效度是相同的, 且 $0\leqslant M_i\leqslant 1$. 事实上 $A_{it}\in[0,1], \sum_{t=1}^{N}Q_t=1, Q_t>0$. 所以 $E(A_i)\in[0,1], \sigma(A_i)\in[0,1]$, 从而 $M_i\in[0,1]$.

特别地, 若对 m 种单项预测方法的预测精度的加权系数先验信息不确知时, 可取 $Q_t=1/N, t=1,2,\cdots,N$.

设 A_t 为第 t 时刻的组合预测精度, e_t 为第 t 时刻的组合预测相对误差, 则有

$$\begin{aligned}A_t&=1-|e_t|=1-|(x_t-\hat{x}_t)/x_t|=1-\left|\sum_{i=1}^{m}l_i\left(x_t-x_{it}\right)/x_t\right|\\&=1-\left|\sum_{i=1}^{m}l_ie_{it}\right|,\quad t=1,2,\cdots,N.\end{aligned}$$

根据定义 5.5.1, 二阶组合预测有效度 M 为

$$\begin{aligned}M=&E(A)(1-\sigma(A))\\=&\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m}l_ie_{it}\right|\right)\left\{1-\left\{\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m}l_ie_{it}\right|\right)^2\right.\right.\\&\left.\left.-\left[\sum_{t=1}^{N}Q_t\left(1-\left|\sum_{i=1}^{m}l_ie_{it}\right|\right)\right]^2\right\}^{\frac{1}{2}}\right\},\end{aligned}$$

其中 $E(A)$ 为组合预测方法的预测精度序列的数学期望, $\sigma(A)$ 为预测精度序列的标准差.

同理 $0 \leqslant M \leqslant 1$. 显然, $E(A)$ 和 $\sigma(A)$ 为各种单项预测方法的加权系数 $l_1, l_2, \cdots, l_m$ 的函数, 所以 M 也为 $l_1, l_2, \cdots, l_m$ 的函数. 记为 $M(l_1, l_2, \cdots, l_m)$.

一般, $M(l_1, l_2, \cdots, l_m)$ 越大表示组合预测方法越有效. 因此, 以二阶组合预测有效度为准则的最优组合预测模型可表示成如下模型

$$\max M = \sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i e_{it}\right|\right) \left\{1 - \left\{\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i e_{it}\right|\right)^2 - \left[\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i e_{it}\right|\right)\right]^2\right\}^{\frac{1}{2}}\right\},$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i = 1, \\ l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{5.5.1}$$

记 $M_{\min} = \min\{M_i, i = 1, 2, \cdots, m\}, M_{\max} = \max\{M_i, i = 1, 2, \cdots, m\}$, 即 $M_{\min}$ 表示 m 种预测方法最小预测有效度, $M_{\max}$ 表示 m 种预测方法最大预测有效度, 则有如下定义.

定义 5.5.2 设 M 为二阶组合预测有效度, 若 $M_{\min} \leqslant M \leqslant M_{\max}$, 则称组合预测模型 (5.5.1) 为非劣性组合预测, 若 $M > M_{\max}$, 则称组合预测模型 (5.5.1) 为优性组合预测.

由于 M 越大表示组合预测方法越有效, 定义 5.5.2 表明, 只有二阶组合预测有效度大于各单项预测有效度中最大者, 则该组合预测模型才为优性组合预测.

定义 5.5.3 若第 i 种预测方法与第 k 种预测方法的预测精度序列的数学期望 $E(A_i), E(A_k)$, 协方差满足以下关系式

$$E(A_i) > E(A_k), \quad \operatorname{cov}(A_i, A_j) < \operatorname{cov}(A_k, A_j), \quad j = 1, 2, \cdots, m,$$

则称第 i 种单项预测方法优超第 k 种单项预测方法. 其中,

$$\operatorname{cov}(A_i, A_j) = \sum_{t=1}^{N} Q_t \left[A_{it} - E(A_i)\right] \left[A_{jt} - E(A_j)\right]$$

为第 i 种预测方法与第 j 种预测方法的预测精度序列的协方差.

特别地, $\operatorname{cov}(A_i, A_i) = \sigma^2(A_i), \operatorname{cov}(A_k, A_k) = \sigma^2(A_k), \sigma^2(A_i), \sigma^2(A_k)$ 分别表示第 i 种预测方法、第 k 种预测方法的预测精度序列的方差.

实际上, 定义 5.5.3 已缊涵了 $\sigma^2(A_i) < \sigma^2(A_k)$. 它表明第 i 种单项预测方法在各个时刻平均预测精度大于第 k 种单项预测方法平均预测精度, 而且第 i 种预测方

法的预测精度序列的方差及与其他 $m-1$ 单项预测方法协方差小于第 k 种单项预测方法的方差及协方差. 就考虑预测精度标准差的预测有效度而言, 直观上可以认为第 i 种单项预测方法要"好于"第 k 种单项预测方法.

5.5.2 非劣性组合预测和优性组合预测存在的充分条件[74]

定理 5.5.1 组合预测模型 (5.5.1) 组合预测方法的预测精度序列 $\{A_t, t=1,2,\cdots,N\}$ 的数学期望 $E(A)$ 和标准差 $\sigma(A)$ 满足如下不等式

$$\sum_{i=1}^{m} l_i E(A_i) \leqslant E(A) \leqslant 1, \tag{5.5.2}$$

$$\sigma(A) \leqslant \sum_{i=1}^{m} l_i \sigma(A_i). \tag{5.5.3}$$

证明 设 $e_{it}=(x_t-x_{it})/x_t$ 为第 i 种单项预测方法的预测相对误差, $i=1,2,\cdots,m, t=1,2,\cdots,N$. 因为

$$0 \leqslant |(x_t-\hat{x}_t)/x_t| = \left|\sum_{i=1}^{m} l_i(x_t-x_{it})/x_t\right| = \left|\sum_{i=1}^{m} l_i e_{it}\right| \leqslant \sum_{i=1}^{m} l_i\,|e_{it}|, \quad t=1,2,\cdots,N.$$

所以 $1 \geqslant A_t = 1-\left|\sum_{i=1}^{m} l_i e_{it}\right| \geqslant 1-\sum_{i=1}^{m} l_i\,|e_{it}|$, 注意到 $\sum_{i=1}^{m} l_i=1$, 则有

$$1-\sum_{i=1}^{m} l_i\,|e_{it}| = \sum_{i=1}^{m} l_i\,(1-|e_{it}|) = \sum_{i=1}^{m} l_i A_{it},$$

因而有

$$\sum_{i=1}^{m} l_i A_{it} \leqslant A_t \leqslant 1. \tag{5.5.4}$$

另外注意到 $\sum_{t=1}^{N} Q_t=1, Q_t>0$, 由式 (5.5.4) 有

$$\sum_{t=1}^{N} Q_t \sum_{i=1}^{m} l_i A_{it} = \sum_{i=1}^{m} l_i \left(\sum_{t=1}^{N} Q_t A_{it}\right) = \sum_{i=1}^{m} l_i E(A_i) \leqslant \sum_{t=1}^{N} Q_t A_t = E(A) \leqslant \sum_{t=1}^{N} Q_t = 1.$$

即 $\sum_{i=1}^{m} l_i E(A_i) \leqslant E(A) \leqslant 1$, 式 (5.5.2) 得证.

因为式 (5.5.4) 表明 A_t 的上界估计为 1, 下界估计为 $\sum_{i=1}^{m} l_i A_{it}$, 所以存在 $\alpha_0 \in [0,1]$, 使得

$$A_t = \alpha_0 \sum_{i=1}^{m} l_i A_{it} + (1-\alpha_0), \quad t=1,2,\cdots,N.$$

则有

$$
\begin{aligned}
E(A) &= \sum_{t=1}^{N} Q_t A_t = \sum_{t=1}^{N} Q_t \left[\alpha_0 \sum_{i=1}^{m} l_i A_{it} + (1-\alpha_0)\right] \\
&= \alpha_0 \sum_{i=1}^{m} l_i \left(\sum_{t=1}^{N} Q_t A_{it}\right) + (1-\alpha_0) \sum_{t=1}^{N} Q_t \\
&= \alpha_0 \sum_{i=1}^{m} l_i E(A_i) + (1-\alpha_0),
\end{aligned} \tag{5.5.5}
$$

从而

$$
\begin{aligned}
\sigma(A) &= \left[\sum_{t=1}^{N} Q_t (A_t - E(A))^2\right]^{\frac{1}{2}} \\
&= \left\{\sum_{t=1}^{N} Q_t \left\{\left[\alpha_0 \sum_{i=1}^{m} l_i A_{it} + (1-\alpha_0)\right] - \left[\alpha_0 \sum_{i=1}^{m} l_i E(A_i) + (1-\alpha_0)\right]\right\}^2\right\}^{\frac{1}{2}} \\
&= \alpha_0 \left\{\sum_{t=1}^{N} Q_t \left[\sum_{i=1}^{m} l_i (A_{it} - E(A_i)\right]^2\right\}^{\frac{1}{2}} \\
&= \alpha_0 \left\{\sum_{i=1}^{m} l_i^2 \sum_{t=1}^{N} Q_t \left[A_{it} - E(A_i)\right]^2\right. \\
&\qquad \left. + \sum_{i \neq j} l_i l_j \sum_{t=1}^{N} Q_t \left[A_{it} - E(A_i)\right]\left[A_{jt} - E(A_j)\right]\right\}^{\frac{1}{2}} \\
&= \alpha_0 \left[\sum_{i=1}^{m} l_i^2 \sigma^2(A_i) + \sum_{i \neq j} l_i l_j \rho_{ij} \sigma(A_i) \sigma(A_j)\right]^{\frac{1}{2}},
\end{aligned} \tag{5.5.6}
$$

其中 $\rho_{ij} = \sum_{t=1}^{N} Q_t \left[A_{it} - E(A_i)\right]\left[A_{jt} - E(A_j)\right] / [\sigma(A_i)\sigma(A_j)]$, 即 ρ_{ij} 为第 i 种单项预测方法的预测精度与第 j 种单项预测方法的预测精度的相关系数.

因为 $\rho_{ij} \in [-1, 1]$, 所以有

$$
\sum_{i=1}^{m} l_i^2 \sigma^2(A_i) + \sum_{i \neq j} l_i l_j \rho_{ij} \sigma(A_i) \sigma(A_j) \leqslant \left[\sum_{i=1}^{m} l_i \sigma(A_i)\right]^2. \tag{5.5.7}
$$

因为 $\alpha_0 \in [0, 1]$, 所以由式 (5.5.6)、(5.5.7) 知

$$
\sigma(A) \leqslant \sum_{i=1}^{m} l_i \sigma(A_i),
$$

从而式 (5.5.3) 成立. 证毕.

定理 5.5.1 实际上分别给出组合预测方法的预测精度序列 $\{A_t, t=1,2,\cdots,N\}$ 的数学期望 $E(A)$ 和标准差 $\sigma(A)$ 的一个估计. 该结论表明, 组合预测方法的预测精度序列数学期望大于或等于各单项预测模型的预测精度序列数学期望的一个线性组合; 组合预测方法的预测精度序列标准差小于或等于各单项预测模型的预测精度序列标准差的一个线性组合. 线性组合系数就是各单项预测模型在组合预测中的加权系数.

定理 5.5.2 组合预测模型 (5.5.1) 的最优解对应的组合预测至少是非劣性组合预测.

证明 设 $\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$为组合预测模型 (5.5.1)的最优解, $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为其任一个可行解, 则目标函数值满足

$$M(l_1^*,l_2^*,\cdots,l_m^*)\geqslant M(l_1,l_2,\cdots,l_m). \tag{5.5.8}$$

由式 (5.5.2)、(5.5.3) 知

$$M(l_1,l_2,\cdots,l_m)=E(A)(1-\sigma(A))\geqslant\sum_{i=1}^m l_iE(A_i)\left(1-\sum_{i=1}^m l_i\sigma(A_i)\right). \tag{5.5.9}$$

显然, $(1,0,\cdots,0)^{\mathrm{T}},(0,1,\cdots,0)^{\mathrm{T}},\cdots,(0,0,\cdots,1)^{\mathrm{T}}$ 这 m 个 m 维列向量分别是组合预测模型 (5.5.1) 的可行解, 由式 (5.5.8)、(5.5.9) 知

$$\begin{cases} M(l_1^*,l_2^*,\cdots,l_m^*)\geqslant M(1,0,\cdots,0)\geqslant E(A_1)(1-\sigma(A_1))=M_1,\\ M(l_1^*,l_2^*,\cdots,l_m^*)\geqslant M(0,1,\cdots,0)\geqslant E(A_2)(1-\sigma(A_2))=M_2,\\ \cdots\cdots\cdots\cdots\\ M(l_1^*,l_2^*,\cdots,l_m^*)\geqslant M(0,0,\cdots,1)\geqslant E(A_m)(1-\sigma(A_m))=M_m,\end{cases}$$

所以

$$\begin{aligned}M(l_1^*,l_2^*,\cdots,l_m^*)&\geqslant \max\{M_i,i=1,2,\cdots,m\}=M_{\max}\\ &\geqslant M_{\min}=\min\{M_i,i=1,2,\cdots,m\}.\end{aligned} \tag{5.5.10}$$

再由定义 5.5.2 得结论成立. 证毕.

定理 5.5.2 表明, 从组合预测有效度角度而言, 组合预测模型 (5.5.1) 的最优解所对应的组合预测均不会比 “最差” 的单项预测方法还要差.

推论 5.5.1 若第 i 种单项预测方法的预测精度序列与第 j 种单项预测方法的预测精度序列的相关系数 $\rho_{ij}\in(-1,1)$, 则组合预测模型 (5.5.1) 最优解所对应的组合预测方法是优性组合预测方法.

证明 因为 $\rho_{ij} \in (-1,1)$, 由式 (5.5.6)、(5.5.7) 知

$$\sigma(A) < \sum_{i=1}^{m} l_i \sigma(A_i).$$

所以式 (5.5.9) 的严格不等号成立, 由定理 5.5.2 证明过程知, (5.5.10) 式也严格不等号成立. 即

$$M(l_1^*, l_2^*, \cdots, l_m^*) > \max\{M_i, i=1,2,\cdots,m\} = M_{\max},$$

由定义 5.5.2 知结论成立. 证毕.

推论 5.5.1 的条件一般来说是很容易满足的, 组合预测模型 (5.5.1) 最优解所对应的组合预测方法要 "好于" 单项预测方法中的 "最好" 者, 可见, 组合预测模型 (5.5.1) 确实综合利用了各单项预测方法提供的信息.

推论 5.5.2 若 $E(A_1) = \min\limits_{1 \leqslant i \leqslant m} \{E(A_i)\}, \sigma(A_1) = \max\limits_{1 \leqslant i \leqslant m} \{\sigma(A_i)\}$, 则简单平均组合预测方法至少为非劣性组合预测方法.

证明 因为 $E(A_1) = \min\limits_{1 \leqslant i \leqslant m} \{E(A_i)\}, \sigma(A_1) = \max\limits_{1 \leqslant i \leqslant m} \{\sigma(A_i)\}$, 所以

$$M_1 = \min\{M_i, i=1,2,\cdots,m\}.$$

在简单平均组合预测方法中, 即令 $l_1 = l_2 = \cdots = l_m = 1/m$, 将其代入 (5.5.9) 式得

$$\begin{aligned} M(1/m, 1/m, \cdots, 1/m) &\geqslant \frac{1}{m}\sum_{i=1}^{m} E(A_i)\left(1 - \frac{1}{m}\sum_{i=1}^{m}\sigma(A_i)\right) \geqslant E(A_1)(1-\sigma(A_1)) \\ &= \min\{M_i, i=1,2,\cdots,m\}. \end{aligned}$$

由定义 5.5.2 知结论成立. 证毕.

5.5.3 冗余信息的判定定理[75]

对于组合预测模型 (5.5.1), 有如下结论成立.

定理 5.5.3 组合预测模型 (5.5.1) 的最优目标函数值是参与组合预测的各单项预测方法总个数 m 的单调不减函数, 即

$$M^*(l_1, l_2, \cdots, l_m) \leqslant M^*(l_1, l_2, \cdots, l_{m+1}),$$

其中 $M^*(l_1, l_2, \cdots, l_m)$ 表示 m 个单项预测方法参与的组合预测模型 (5.5.1) 对应的最大组合预测有效度, $M^*(l_1, l_2, \cdots, l_{m+1})$ 表示再增加一个单项预测方法参与的组合预测模型对应的最大组合预测有效度.

证明 设 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$ 为 m 个单项预测方法参与的组合预测模型 (5.5.1) 的最优解, 由式 (5.5.1) 知

$$
\begin{aligned}
&M^*(l_1, l_2, \cdots, l_m) \\
&= \sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i^* e_{it}\right|\right) \left\{1 - \left\{\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i^* e_{it}\right|\right)^2 \right.\right. \\
&\quad \left.\left. - \left[\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i^* e_{it}\right|\right)\right]^2 \right\}^{\frac{1}{2}} \right\},
\end{aligned}
$$

其中 $\sum\limits_{i=1}^{m} l_i^* = 1, l_i^* \geqslant 0, i = 1, 2, \cdots, m.$

同理, 设 $\bar{\boldsymbol{L}} = (\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1})^{\mathrm{T}}$ 为再增加一个单项预测方法, 共 $m+1$ 个单项预测方法参与的组合预测模型 (5.5.1) 的最优解, 则有

$$
\begin{aligned}
&M^*(l_1, l_2, \cdots, l_m, l_{m+1}) \\
&= \sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m+1} \bar{l}_i e_{it}\right|\right) \left\{1 - \left\{\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m+1} \bar{l}_i e_{it}\right|\right)^2 \right.\right. \\
&\quad \left.\left. - \left[\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m+1} \bar{l}_i e_{it}\right|\right)\right]^2 \right\}^{\frac{1}{2}} \right\},
\end{aligned}
$$

其中 $\sum\limits_{i=1}^{m+1} \bar{l}_i = 1, \bar{l}_i \geqslant 0, i = 1, 2, \cdots, m+1.$

令 $\boldsymbol{L} = (l_1^*, l_2^*, \cdots, l_m^*, 0)^{\mathrm{T}}$, 显然, $\boldsymbol{L}$ 为 $m+1$ 个单项预测方法参与的组合预测模型 (5.5.1) 的可行解, 则有

$$
\begin{aligned}
&M^*(l_1, l_2, \cdots, l_m, l_{m+1}) \geqslant M(l_1^*, l_2^*, \cdots, l_m^*, 0) \\
&= \sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i^* e_{it}\right|\right) \left\{1 - \left\{\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i^* e_{it}\right|\right)^2 \right.\right. \\
&\quad \left.\left. - \left[\sum_{t=1}^{N} Q_t \left(1 - \left|\sum_{i=1}^{m} l_i^* e_{it}\right|\right)\right]^2 \right\}^{\frac{1}{2}} \right\} \\
&= M^*(l_1, l_2, \cdots, l_m).
\end{aligned}
$$

从而结论成立. 证毕.

定理 5.5.3 表明, 组合预测模型 (5.5.1) 可能存在冗余预测方法. 通常认为组合预测模型 (5.5.1) 的最大组合预测有效度, 一定是参与组合预测的单项预测方法个

数 m 的严格单调增加函数, 然而, 定理 5.5.3 证明了当再增加一个单项预测方法时, 组合预测模型 (5.5.1) 的最大组合预测有效度可能不变.

下面一个定理为冗余信息提供了判定方法. 为此先证明一个引理.

引理 5.5.1 设 $\boldsymbol{A}=(A_{it})_{m\times N}$ 为组合预测模型 (5.5.1) 的预测精度矩阵, $\boldsymbol{A}$ 的秩为 1, 若 $E(A_1)=\max\{E(A_i),i=1,2,\cdots,m\}$, 则有

$$\rho_{ij}=1,\quad i,j=1,\cdots,m,i\neq j,\quad M_1=\max_{1\leqslant i\leqslant m}\{M_i\},$$

其中 ρ_{ij} 为第 i 种单项预测方法的预测精度序列与第 j 种单项预测方法的预测精度序列的相关系数, M_i 为第 i 种单项预测方法的预测有效度.

证明 对预测精度矩阵 $\boldsymbol{A}$ 进行分块, 设 $\boldsymbol{A}=(\boldsymbol{a}_1^{\mathrm{T}},\boldsymbol{a}_2^{\mathrm{T}},\cdots,\boldsymbol{a}_m^{\mathrm{T}})^{\mathrm{T}}$, 即 $\boldsymbol{a}_i=(A_{i1},A_{i2},\cdots,A_{iN})$ 为 $\boldsymbol{A}$ 的第 i 行, $i=1,2,\cdots,m$. 先证明 $\boldsymbol{a}_1\neq\boldsymbol{0}$.

若 $\boldsymbol{a}_1=\boldsymbol{0}$, 则有 $E(A_1)=\sum_{t=1}^{N}Q_tA_{1t}=0$, 因为 $E(A_1)=\max\limits_{1\leqslant i\leqslant m}\{E(A_i)\}$, 注意到 $\boldsymbol{A}$ 为非负矩阵, 所以 $E(A_i)=0$, 从而 $\boldsymbol{a}_i=\boldsymbol{0},i=2,3,\cdots,m$, 即 $\boldsymbol{A}=\boldsymbol{0}$, 这与 $\boldsymbol{A}$ 的秩为 1 矛盾. 所以 $\boldsymbol{a}_1\neq\boldsymbol{0}$ 成立.

因为 $\boldsymbol{A}$ 的秩为 1, 所以存在比例常数 $k_i\neq 0$, 使得 $\boldsymbol{a}_i=k_i\boldsymbol{a}_1,i=2,3,\cdots,m$, 则有

$$E(A_i)=\sum_{t=1}^{N}Q_tA_{it}=\sum_{t=1}^{N}Q_tk_iA_{1t}=k_iE(A_1),\quad i=2,\cdots,m,\tag{5.5.11}$$

$$\begin{aligned}\sigma(A_i)&=\left[\sum_{t=1}^{N}Q_t(A_{it}-E(A_i))^2\right]^{\frac{1}{2}}\\&=\left\{\sum_{t=1}^{N}Q_t[k_iA_{1t}-k_iE(A_1)]^2\right\}^{\frac{1}{2}}=k_i\sigma(A_1).\end{aligned}\tag{5.5.12}$$

所以

$$\rho_{ij}=\sum_{t=1}^{N}Q_t(A_{it}-E(A_i))(A_{jt}-E(A_j))/(\sigma(A_i)\sigma(A_j))=1.\tag{5.5.13}$$

又由式 (5.5.11) 及 $E(A_1)=\max\limits_{1\leqslant i\leqslant m}\{E(A_i)\}$ 知

$$k_i\in[0,1].\tag{5.5.14}$$

由定义 5.5.1 和式 (5.5.11), (5.5.12) 知

$$M_i=E(A_i)(1-\sigma(A_i))=k_iE(A_1)(1-k_i\sigma(A_1)).$$

M_i 可视为 k_i 的一元二次函数, 在 $k_i \in [0,1]$ 区间上 M_i 为 k_i 的严格单调增函数. 所以 $M_i \leqslant E(A_1)(1-\sigma(A_1)) = M_1$. 证毕.

定理 5.5.4 若组合预测模型的相对误差矩阵 $\boldsymbol{E} = (e_{it})_{m\times N}$ 中, 任意第 t 列元素 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的符号完全相同, $t = 1,2,\cdots,N$, 且满足上述引理 5.5.1 的条件, 则组合预测模型 (5.5.1) 的冗余度为 $(m-1)/m$. 即后面的 $m-1$ 种单项预测方法为冗余预测方法.

证明 因为 $\boldsymbol{E}$ 中任意第 t 列元素 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的符号完全相同, $t \in \{1,2,\cdots, N\}$.

所以 $\left|\sum\limits_{i=1}^{m} l_i e_{it}\right| = \sum\limits_{i=1}^{m} l_i |e_{it}|, t \in \{1,2,\cdots,N\}$, 注意到 $\sum\limits_{i=1}^{m} l_i = 1, l_i \geqslant 0$, 于是有

$$A_t = 1 - \left|\sum_{i=1}^{m} l_i e_{it}\right| = 1 - \sum_{i=1}^{m} l_i |e_{it}| = \sum_{i=1}^{m} l_i (1 - |e_{it}|) = \sum_{i=1}^{m} l_i A_{it}, \tag{5.5.15}$$

$$E(A) = \sum_{t=1}^{N} Q_t A_t = \sum_{t=1}^{N} Q_t \sum_{i=1}^{m} l_i A_{it} = \sum_{i=1}^{m} l_i \left(\sum_{t=1}^{N} Q_t A_{it}\right) = \sum_{i=1}^{m} l_i E(A_i). \tag{5.5.16}$$

所以

$$\begin{aligned}\sigma(A) &= \left[\sum_{t=1}^{N} Q_t (A_t - E(A))^2\right]^{\frac{1}{2}} = \left\{\sum_{t=1}^{N} Q_t \left[\sum_{i=1}^{m} l_i A_{it} - \sum_{i=1}^{m} l_i E(A_i)\right]^2\right\}^{\frac{1}{2}} \\ &= \left\{\sum_{i=1}^{m} l_i^2 \sigma^2(A_i) + \sum_{i\neq j} l_i l_j \rho_{ij} \sigma(A_i)\sigma(A_j)\right\}^{\frac{1}{2}},\end{aligned} \tag{5.5.17}$$

其中 ρ_{ij} 为第 i 种单项预测方法的预测精度与第 j 种单项预测方法的预测精度的相关系数.

因为上述引理的条件满足, 所以有 $\rho_{ij} = 1$, 即

$$\sigma(A) = \sum_{i=1}^{m} l_i \sigma(A_i). \tag{5.5.18}$$

由定义 5.5.1 及式 (5.5.11)、(5.5.12)、(5.5.14)、(5.5.16)、(5.5.18) 知

$$\begin{aligned}M(l_1, l_2, \cdots, l_m) &= E(A)(1-\sigma(A)) = \sum_{i=1}^{m} l_i E(A_i) \left(1 - \sum_{i=1}^{m} l_i \sigma(A_i)\right) \\ &= \sum_{i=1}^{m} l_i k_i E(A_1) \left(1 - \sum_{i=1}^{m} l_i k_i \sigma(A_1)\right), \quad 令 k_1 = 1.\end{aligned} \tag{5.5.19}$$

因为 $\sum\limits_{i=1}^{m} l_i = 1, k_i \in (0,1]$, 所以 $0 < \sum\limits_{i=1}^{m} l_i k_i \leqslant \max\{k_i, i = 1,2,\cdots,m\} = k_1 = 1$.

同理, $M(l_1,l_2,\cdots,l_m)$ 可视为 $\sum\limits_{i=1}^{m} l_ik_i$ 的一元二次函数, 它在 $[0,1]$ 区间上为 $\sum\limits_{i=1}^{m} l_ik_i$ 的严格单调增函数. 所以有

$$M(l_1,l_2,\cdots,l_m) \leqslant E(A_1)(1-\sigma(A_1)) = M_1.$$

由式 (5.5.19) 知

$$M(1,0,\cdots,0) = M_1.$$

所以 $\boldsymbol{L}=(1,0,\cdots,0)^{\mathrm{T}}$ 为组合预测模型 (5.5.1) 的最优解. 即后面的 $m-1$ 种单项预测方法为冗余预测方法.

若组合预测模型的冗余度 $(m-1)/m$, 它表示组合预测中只有一个单项预测方法提供有效信息, 因此我们称这种组合预测为无效组合预测. 上述定理给出无效组合预测存在的一种情形.

定理 5.5.5 在组合预测模型 (5.5.1) 中, 若第 i 种单项预测方法优超第 k 种单项预测方法, 则组合预测模型 (5.5.1) 的冗余度至少为 $1/m$.

证明 令 $\boldsymbol{V}=(v_{ij})_{m\times m}$ 为 m 种单项预测方法的预测精度序列的协方差矩阵, 其中 $v_{ii}=\sigma^2(A_i)$ 表示第 i 种预测方法的预测精度序列的方差, $v_{ij}=\mathrm{cov}(A_i,A_j)$ 为第 i 种预测方法与第 j 种预测方法的预测精度序列的协方差, 显然 $\boldsymbol{V}=(v_{ij})_{m\times m}$ 为 m 阶对称矩阵, 由式 (5.5.5) 和式 (5.5.6) 知, 存在 $\alpha_0\in[0,1]$, 满足

$$E(A)=\alpha_0\sum_{i=1}^{m} l_iE(A_i)+(1-\alpha_0),$$

$$\sigma^2(A)=\alpha_0^2\left[\sum_{i=1}^{m} l_i^2\sigma^2(A_i)+\sum_{i\neq j} l_il_j\rho_{ij}\sigma(A_i)\sigma(A_j)\right]$$

注意到 $\rho_{ij}\sigma(A_i)\sigma(A_j)=\mathrm{cov}(A_i,A_j)=v_{ij}$, 则有

$$\sigma^2(A)=\alpha_0^2\left[\sum_{i=1}^{m} l_i^2\sigma^2(A_i)+\sum_{i\neq j} l_il_j\rho_{ij}\sigma(A_i)\sigma(A_j)\right]=\alpha_0^2\boldsymbol{L}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{L},$$

其中 $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为组合预测加权向量.

假设 $\boldsymbol{L}^*=(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型 (5.5.1) 的最优解, 且第 k 种单项预测方法不为冗余预测方法, 即 $l_k^*>0$, 且 $\sum\limits_{i=1}^{m} l_i^*=1$. 则与 $\boldsymbol{L}^*$ 对应组合预测模型 (5.5.1) 的目标函数值为

$$M(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)=\left[\alpha_0\sum_{j=1}^{m} l_j^*E(A_j)+(1-\alpha_0)\right]\left[1-\alpha_0\left(\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{V}\boldsymbol{L}^*\right)^{\frac{1}{2}}\right].$$

构造向量

$$\hat{\boldsymbol{L}} = \boldsymbol{L}^* + \tilde{\boldsymbol{L}} = (l_1^*, \cdots, l_i^* + l_k^*, \cdots, 0, \cdots, l_m^*)^{\mathrm{T}},$$

其中 $\tilde{\boldsymbol{L}} = (0, \cdots, l_k^*, \cdots, -l_k^*, \cdots, 0)^{\mathrm{T}}$, 即 $\tilde{\boldsymbol{L}}$ 的第 i 个分量为 l_k^*, 第 k 个分量为 $-l_k^*$, 其余 $m-2$ 个分量全为 0. 显然, $\hat{\boldsymbol{L}}$ 为组合预测模型 (5.5.1) 的一可行解, 则与 $\hat{\boldsymbol{L}}$ 对应组合预测模型 (5.5.1) 的目标函数值为

$$\begin{aligned} & M(l_1^*, \cdots, l_i^* + l_k^*, \cdots, 0, \cdots, l_m^*) \\ = & \left\{ \alpha_0 \left[\sum_{j=1}^{m} l_j^* E(A_j) + l_k^* E(A_i) - l_k^* E(A_k) \right] + (1-\alpha_0) \right\} \left[1 - \alpha_0 \left(\hat{\boldsymbol{L}}^{\mathrm{T}} \boldsymbol{V} \hat{\boldsymbol{L}} \right)^{\frac{1}{2}} \right], \end{aligned}$$

$$\begin{aligned} \hat{\boldsymbol{L}}^{\mathrm{T}} \boldsymbol{V} \hat{\boldsymbol{L}} &= (\boldsymbol{L}^* + \tilde{\boldsymbol{L}})^{\mathrm{T}} \boldsymbol{V} (\boldsymbol{L}^* + \tilde{\boldsymbol{L}}) = \boldsymbol{L}^{*\mathrm{T}} \boldsymbol{V} \boldsymbol{L}^* + 2\tilde{\boldsymbol{L}}^{\mathrm{T}} \boldsymbol{V} \boldsymbol{L}^* + \tilde{\boldsymbol{L}}^{\mathrm{T}} \boldsymbol{V} \tilde{\boldsymbol{L}} \\ &= \boldsymbol{L}^{*\mathrm{T}} \boldsymbol{V} \boldsymbol{L}^* + 2 l_k^* \sum_{j \neq k} l_j^* (v_{ij} - v_{kj}) + l_k^{*2} (v_{ii} - v_{kk}). \end{aligned} \tag{5.5.20}$$

又因为第 i 种单项预测方法优超第 k 种单项预测方法, 所以由定义 5.5.3 知

$$E(A_i) > E(A_k), \quad v_{ii} < v_{kk}, \quad v_{ij} < v_{kj}, \quad j = 1, \cdots, m, \quad j \neq k.$$

注意到假设 $l_k^* > 0$, 且由式 (5.5.20) 知

$$\hat{\boldsymbol{L}}^{\mathrm{T}} \boldsymbol{V} \hat{\boldsymbol{L}} < \boldsymbol{L}^{*\mathrm{T}} \boldsymbol{V} \boldsymbol{L}^*, \quad \sum_{j=1}^{m} l_j^* E(A_j) + l_k^* E(A_i) - l_k^* E(A_k) > \sum_{j=1}^{m} l_j^* E(A_j).$$

所以

$$M(l_1^*, \cdots, l_i^* + l_k^*, \cdots, 0, \cdots, l_m^*) > M(l_1^*, \cdots, l_i^*, \cdots, l_k^*, \cdots, l_m^*).$$

而这与 $\boldsymbol{L} = (l_1^*, \cdots, l_i^*, \cdots, l_k^*, \cdots, l_m^*)^{\mathrm{T}}$ 为组合预测模型 (5.5.1) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 此即组合预测模型 (5.5.1) 的冗余度至少为 $1/m$. 证毕.

5.5.4 组合预测模型的近似求解方法[76]

从模型 (5.5.1) 的目标函数可知, 由于绝对值的存在使得模型 (5.5.1) 的目标函数不可导, 因此模型 (5.5.1) 实为不可微非线性规划, 同时组合预测权系数 $l_1, l_2, \cdots, l_m$ 所在的位置也很分散, 所以当 N 和 m 较大时, 模型 (5.5.1) 的计算复杂度较大. 下面讨论组合预测模型 (5.5.1) 的一种近似求解方法. 基本思路是把这种不可微非线性规划转化成可微非线性规划来求解.

根据 (5.5.5)、(5.5.6) 两式知, 组合预测模型 (5.5.1) 等价于如下最优化模型

$$
\begin{aligned}
&\max M(l_1,l_2,\cdots,l_m)\\
&=\left[\alpha_0\sum_{i=1}^{m}l_iE(A_i)+(1-\alpha_0)\right]\\
&\quad\times\left\{1-\alpha_0\left[\sum_{i=1}^{m}l_i^2\sigma^2(A_i)+\sum_{i\neq j}l_il_j\rho_{ij}\sigma(A_i)\sigma(A_j)\right]^{\frac{1}{2}}\right\},\\
&\text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{m}l_i=1,\\ l_i\geqslant 0,i=1,2,\cdots,m,\end{cases}
\end{aligned}
\tag{5.5.21}
$$

其中 $\alpha_0\in[0,1]$ 为常数.

下面介绍最优化模型 (5.5.2) 的特例, 为此先引进一个推论.

推论 5.5.3 组合预测模型的相对误差矩阵 $\boldsymbol{E}=(e_{it})_{m\times N}$ 中, 若任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t=1,2,\cdots,N$, 则模型 (5.5.1) 的组合预测精度序列 $\{A_t,t=1,2,\cdots,N\}$ 的数学期望 $E(A)$ 和标准差 $\sigma(A)$ 满足

$$
\begin{cases}
E(A)=\displaystyle\sum_{i=1}^{m}l_iE(A_i),\\
\sigma(A)=\left[\displaystyle\sum_{i=1}^{m}l_i^2\sigma^2(A_i)+\sum_{i\neq j}l_il_j\rho_{ij}\sigma(A_i)\sigma(A_j)\right]^{\frac{1}{2}},
\end{cases}
\tag{5.5.22}
$$

其中 $\sum_{i=1}^{m}l_i=1,l_i\geqslant 0,i=1,2,\cdots,m,E(A_i)$ 为第 i 种预测方法的预测精度序列的数学期望, $\sigma(A_i)$ 为第 i 种预测方法的预测精度序列的标准差. e_{it} 为第 i 种预测方法第 t 时刻的预测相对误差, $i=1,2,\cdots,m,t=1,2,\cdots,N$.

证明参见定理 5.5.4 的证明过程.

推论 5.5.3 表明, 组合预测模型的相对误差矩阵 $\boldsymbol{E}=(e_{it})_{m\times N}$ 中, 若任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t=1,2,\cdots,N$, 则二阶组合预测有效度 $M(l_1,l_2,\cdots,l_m)$ 为

$$
\begin{aligned}
M(l_1,l_2,\cdots,l_m)&=E(A)[1-\sigma(A)]\\
&=\sum_{i=1}^{m}l_iE(A_i)\left\{1-\left[\sum_{i=1}^{m}l_i^2\sigma^2(A_i)+\sum_{i\neq j}l_il_j\rho_{ij}\sigma(A_i)\sigma(A_j)\right]^{\frac{1}{2}}\right\}.
\end{aligned}
$$

特别地, 若任意两种预测方法的预测精度序列的相关系数为零, 即 $\rho_{ij}=0,i,j=$

$1,2,\cdots,m,i\neq j$. 则最优化模型 (5.5.21) 可进一步简化为如下模型

$$\max M(l_1,l_2,\cdots,l_m)=\sum_{i=1}^{m}l_iE(A_i)\left\{1-\left[\sum_{i=1}^{m}l_i^2\sigma^2(A_i)\right]^{\frac{1}{2}}\right\},$$

$$\text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{m}l_i=1,\\ l_i\geqslant 0,\quad i=1,2,\cdots,m.\end{cases}\tag{5.5.23}$$

模型 (5.5.21)、模型 (5.5.23) 与模型 (5.5.1) 相比一个显著特点是它去掉了模型 (5.5.1) 中目标函数的绝对值, 且模型 (5.5.2)、模型 (5.5.3) 均为可微的非线性规划, 已有现成的计算软件, 易于求解.

另外, 关于模型 (5.5.2) 中 α_0 的选取, 可根据各单项预测方法相对误差在相同时刻 t 的符号的不一致程度取不同的 α_0. 在推论 5.5.3 的条件满足的情况下, 取 $\alpha_0=1$, 若相对误差的符号不一致的程度越严重, 则 α_0 的取值将偏小一点.

5.5.5 实例分析

本节应用文献 [77, 78] 的数据进行实例分析. 某省 1980~1992 年化工行业专门人才数的统计资料及采用两种预测模型的预测数值, 预测精度序列值及相对误差见表 5.5.1.

表 5.5.1 两种预测模型的预测数值、预测精度序列值及相对误差

年份t	实际值	单项预测方法预测值		预测精度序列值		单项预测方法相对误差	
		x_{1t}	x_{2t}	A_{1t}	A_{2t}	e_{1t}	e_{2t}
1	6014	5266.7	4348.8	0.8758	0.7271	0.1242	0.2729
2	6398	6027.1	5393.8	0.9420	0.8431	0.0580	0.1569
3	6181	6888.7	6520.2	0.8855	0.9461	−0.1145	−0.0539
4	7416	7851.6	7727.4	0.9406	0.9580	−0.0594	−0.0420
5	9076	8915.6	9015.6	0.9823	0.9934	0.0177	0.0066
6	8399	10089.9	10385.9	0.7997	0.7636	−0.2003	−0.2364
7	9215	11347.4	11835.2	0.7792	0.7151	−0.2208	−0.2849
8	11125	12715.2	13366.5	0.8571	0.7985	−0.1429	−0.2015
9	14120	14185.2	14978.8	0.9954	0.9392	−0.0046	−0.0608
10	17238	15754	16672.2	0.9139	0.9672	0.0861	0.0328
11	18689	17425.9	18446.5	0.9324	0.9870	0.0676	0.0130
12	20592	19198.5	20301.9	0.9323	0.9859	0.0677	0.0141
13	22665	21072.4	22238.3	0.9297	0.9812	0.0703	0.0188

通过表 5.5.1 的数据不难发现, 这两种单项预测方法在任何某个时刻 t 的相对误差符号均相同, 即对任意 $t\in\{1,2,\cdots,13\}$, e_{1t} 与 e_{2t} 符号均相同, 从而推论 5.5.3

的条件满足，所以由式 (5.5.21)、(5.5.22) 得

$$E(A)=l_1E(A_1)+l_2E(A_2),$$

$$\sigma(A)=\left[l_1^2\sigma^2(A_1)+l_2^2\sigma^2(A_2)+2l_1l_2\rho_{12}\sigma(A_1)\sigma(A_2)\right]^{\frac{1}{2}}.$$

经计算得，两种预测方法的预测精度序列的均值、方差及相关系数分别为

$$E(A_1)=0.9051,\quad \sigma^2(A_1)=0.0038,$$

$$E(A_2)=0.8929,\quad \sigma^2(A_2)=0.0105,\quad \rho_{12}=0.7884.$$

所以将这些数据代入最优化模型 (5.5.21) 中可得如下最优化模型

$$\begin{aligned}&\max M(l_1,l_2)=(0.9051l_1+0.8929l_2)\left[1-(0.0038l_1^2+0.00996l_1l_2+0.0105l_2^2)\right]^{\frac{1}{2}},\\&\text{s.t.}\begin{cases}l_1+l_2=1,\\l_1\geqslant 0,l_2\geqslant 0.\end{cases}\end{aligned}\tag{5.5.24}$$

利用 Matlab 最优化工具箱计算得上述最优化模型的解为

$$l_1^*=1,\quad l_2^*=0,\quad M(l_1^*,l_2^*)=0.8493,$$

其中目标函数 $M(l_1^*,l_2^*)$ 为预测有效度.

显然，最优化模型解的结果表明第二种单项预测方法为冗余预测方法. 另外，从定义 5.5.3 的角度来看上面的计算结果. 比较两种预测方法的预测精度序列的均值和方差结果则有

$$E(A_1)>E(A_2),\quad \sigma^2(A_1)<\sigma^2(A_2).$$

由定义 5.5.3 知，第一种单项预测方法优超第二种单项预测方法，所以本例正好满足定理 5.5.5 的条件，从而由定理 5.5.5 的结论知，第二种单项预测方法为冗余预测方法. 所以模型 (5.5.24) 的最优解的结果与定理 5.5.5 的结论是一致的. 从而实例分析也表明了定理 5.5.5 的结论的正确性.

另外，从定义 5.5.3 来看，本例直观上可以认为第一种单项预测方法要 “好于” 第二种单项预测方法. 第二种单项预测方法只提供冗余信息. 经计算最优化模型 (5.5.24) 的解和定义 5.5.3 的直观具有一致性，这个实例分析从另外一个角度说明了定义 5.5.3 具有一定的道理.

然而文献 [78] 中给出的近似算法中存在一个缺陷. 文献 [78] 认为两种单项预测方法是相互独立的，即认为 $\rho_{12}=0$，从而它的近似最优解 [78] 为

$$l_1^{**}=0.7343,\quad l_2^{**}=0.2657,\quad M(l_1^{**},l_2^{**})=0.8383,$$

其中目标函数 $M(l_1^{**}, l_2^{**})$ 也表示预测有效度.

显然 $M(l_1^*, l_2^*) > M(l_1^{**}, l_2^{**})$. 即以预测有效度作为判断准则, 本节提出的算法结果优于文献 [78] 的算法, 结果令人满意.

本节模型 (5.5.21) 算法的计算步骤可概括为:

(1) 利用 m 个预测方法对指标序列进行预测得 $\{x_{it}, i = 1, 2, \cdots, m, t = 1, 2, \cdots, N\}$, 由此计算不同时刻的精度序列 $\{A_{it}, i = 1, 2, \cdots, m, t = 1, 2, \cdots, N\}$ 及其均值 $E(A_i)$, 方差 $\sigma^2(A_i)$ 及相关系数 $\rho_{ij}, i, j = 1, 2, \cdots, m$.

(2) 给定 $\alpha_0 \in [0, 1]$, 按模型 (5.5.21) 计算最优的权系数 $l_1^*, l_2^*, \cdots, l_m^*$.

(3) 按 $x_t = \sum_{i=1}^{m} l_i^* x_{it}$ 进行组合预测, $t = N+1, N+2, \cdots$.

本节就二阶预测有效度对组合预测模型的优性特征和组合结构作了进一步分析, 得到了一些有益的结论. 并对二阶预测有效度的组合预测模型讨论其近似求解方法, 实例计算结果表明, 该模型的确综合利用了各种单项预测方法所提供的信息.

5.6 回归型组合预测模型的权系数估计及其显著性检验[79]

目前以预测误差平方和达到最小的组合预测权系数的计算主要利用数学规划模型求得. 但在实际问题中, 有时会出现某个单项预测模型在组合预测中的权系数接近于零. 这说明组合预测对该单项预测模型的偏重程度较小. 能否在组合预测中剔除该单项预测模型, 有必要对组合预测的权系数进行显著性检验. 因此, 本节从统计学的角度考虑, 首先建立组合预测的线性模型, 然后给出其权系数的最小二乘 (LS) 估计及其统计假设检验.

5.6.1 组合预测线性模型的建立

设对同一预测对象的某个指标序列为 $\{x_t, t = 1, 2, \cdots, N\}$, 存在 m 种单项预测方法对其进行预测, 设 x_{it} 为第 i 种单项预测方法在第 t 时刻的预测值 $i = 1, 2, \cdots, m, t = 1, 2, \cdots, N$,

$$\hat{x}_t = l_1 x_{1t} + l_2 x_{2t} + \cdots + l_m x_{mt}, \tag{5.6.1}$$

则 $\hat{x}_t$ 为 x_t 的组合预测值, $l_1, l_2, \cdots, l_m$ 分别为 m 种单项预测方法的加权系数, 令

$$e_t = x_t - \hat{x}_t, \tag{5.6.2}$$

则 e_t 为组合预测在第 t 时刻的预测误差, e_t 为一随机变量, 设 e_t 服从均值为零, 方

差为 σ^2 的正态分布. 由 (5.6.1)、(5.6.2) 两式可得如下线性模型

$$\begin{cases} x_1 = l_1 x_{1t} + l_2 x_{2t} + \cdots + l_m x_{mt} + e_1, \\ x_2 = l_1 x_{1t} + l_2 x_{2t} + \cdots + l_m x_{mt} + e_2, \\ \qquad \cdots\cdots\cdots\cdots \\ x_N = l_1 x_{1t} + l_2 x_{2t} + \cdots + l_m x_{mt} + e_N. \end{cases} \tag{5.6.3}$$

显然, 加权系数应满足

$$l_1 + l_2 + \cdots + l_m = 1. \tag{5.6.4}$$

令 $\boldsymbol{Y} = (x_1 \quad x_2 \quad \cdots \quad x_N)^{\mathrm{T}}, \boldsymbol{e} = (e_1 \quad e_2 \quad \cdots \quad e_N)^{\mathrm{T}}, \boldsymbol{L} = (l_1 \quad l_2 \quad \cdots \quad l_m)^{\mathrm{T}}, \boldsymbol{R} = (1 \quad 1 \quad \cdots \quad 1)^{\mathrm{T}}$, $\boldsymbol{R}$ 是元素全为 1 的 m 维列向量,

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{m1} \\ x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & & \vdots \\ x_{1N} & x_{2N} & \cdots & x_{mN} \end{bmatrix}.$$

则式 (5.6.3)、(5.6.4) 可化为如下矩阵形式

$$\begin{cases} \boldsymbol{Y} = \boldsymbol{X}\boldsymbol{L} + \boldsymbol{e}, \\ \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1. \end{cases} \tag{5.6.5}$$

此为含等式约束的线性模型.

5.6.2 含等式约束的组合预测线性模型的权系数 LS 估计及其性质

先介绍含等式约束的 LS 估计的定义.

定义 5.6.1 对于线性模型 (5.6.5), 若 $\hat{\boldsymbol{L}}$ 满足

$$\|\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{L}}\|^2 = \min_{\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}=1} \|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{L}\|^2,$$

则称 $\hat{\boldsymbol{L}}$ 为权系数向量的 $\boldsymbol{L}$ 的一个 LS 估计, 其中 $\|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{L}\|^2 = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{L})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{L})$.

为了求组合预测权系数向量的 $\boldsymbol{L}$ 的一个 LS 估计, 引进下述一个有关矩阵的事实.

引理 5.6.1[49] 设 $\boldsymbol{S}_{p\times p}$ 为对称可逆矩阵, $\boldsymbol{L}_{q\times p}$ 为行满秩矩阵, 则有

$$\begin{pmatrix} \boldsymbol{S} & \boldsymbol{L}^{\mathrm{T}} \\ \boldsymbol{L} & \boldsymbol{0} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{S}^{-1} - \boldsymbol{S}^{-1}\boldsymbol{L}^{\mathrm{T}}(\boldsymbol{L}\boldsymbol{S}^{-1}\boldsymbol{L}^{\mathrm{T}})^{-1}\boldsymbol{L}\boldsymbol{S}^{-1} & \boldsymbol{S}^{-1}\boldsymbol{L}^{\mathrm{T}}(\boldsymbol{L}\boldsymbol{S}^{-1}\boldsymbol{L}^{\mathrm{T}})^{-1} \\ (\boldsymbol{L}\boldsymbol{S}^{-1}\boldsymbol{L}^{\mathrm{T}})^{-1}\boldsymbol{L}\boldsymbol{S}^{-1} & -(\boldsymbol{L}\boldsymbol{S}^{-1}\boldsymbol{L}^{\mathrm{T}})^{-1} \end{pmatrix}.$$

在线性模型 (5.6.5) 中, $\boldsymbol{X}$ 为 $N \times n$ 矩阵, 一般有 $N > n$, 且 $\boldsymbol{X}$ 为列满秩矩阵, 则有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 为对称可逆矩阵.

定理 5.6.1　线性模型 (5.6.5) 的 LS 估计为

$$\hat{\boldsymbol{L}} = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{12},$$

其中

$$\boldsymbol{A}_{11} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} - (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R})^{-1}\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1},$$

$$\boldsymbol{A}_{12} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R})^{-1}.$$

证明　组合预测线性模型 (5.6.5) 的 LS 估计即求如下优化问题的最优解

$$\begin{aligned}&\min \|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{L}\|^2,\\ &\text{s.t.}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1.\end{aligned}$$

由 Lagrange 乘子法, 构造辅助函数

$$Q(\boldsymbol{L}, \lambda) = \|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{L}\|^2 + 2\lambda(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} - 1).$$

由极值得当必要条件可得如下正规方程组

$$\begin{cases}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L} + \lambda\boldsymbol{R} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y},\\ \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1.\end{cases} \tag{5.6.6}$$

由行列式的降阶定理知, 式 (5.6.6) 方程组系数矩阵的行列式非零, 所以式 (5.6.6) 方程组存在唯一解, 由引理 5.6.1 知

$$\begin{pmatrix}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} & \boldsymbol{R}\\ \boldsymbol{R}^{\mathrm{T}} & \boldsymbol{0}\end{pmatrix}^{-1} = \begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix},$$

其中 $\boldsymbol{A}_{11}, \boldsymbol{A}_{12}$ 同上, $\boldsymbol{A}_{21} = (\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R})^{-1}\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}, \boldsymbol{A}_{22} = (\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R})^{-1}$, 所以 $\hat{\boldsymbol{L}} = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{12}, \lambda = \boldsymbol{A}_{21}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{22}$, 下证 $\hat{\boldsymbol{L}}$ 为优化问题的最优解.

设 $\boldsymbol{\beta}$ 为满足 $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{\beta} = 1$ 的 n 维列向量, 则

$$\begin{aligned}\|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 &= \|(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{L}}) + (\boldsymbol{X}\hat{\boldsymbol{L}} - \boldsymbol{X}\boldsymbol{\beta})\|^2\\ &= \|\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{L}}\|^2 + \|\boldsymbol{X}\hat{\boldsymbol{L}} - \boldsymbol{X}\boldsymbol{\beta}\|^2 + 2(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{L}})^{\mathrm{T}}(\boldsymbol{X}\hat{\boldsymbol{L}} - \boldsymbol{X}\boldsymbol{\beta}),\end{aligned}$$

其中

$$(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{L}})^{\mathrm{T}}(\boldsymbol{X}\hat{\boldsymbol{L}} - \boldsymbol{X}\boldsymbol{\beta}) = \lambda\boldsymbol{R}^{\mathrm{T}}(\hat{\boldsymbol{L}} - \boldsymbol{\beta}) = \lambda(\boldsymbol{R}^{\mathrm{T}}\hat{\boldsymbol{L}} - \boldsymbol{R}^{\mathrm{T}}\boldsymbol{\beta}) = 0.$$

所以 $\|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 \geqslant \|(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{L}})\|^2$, 即 $\hat{\boldsymbol{L}} = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{12}$ 为式 (5.6.5) 的 LS 估计. 证毕.

定理 5.6.2 $\hat{\boldsymbol{L}} = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{12}$ 为线性模型 (5.6.5) 中 $\boldsymbol{L}$ 的最佳线性无偏估计.

证明 因为 $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{L} + \boldsymbol{e}$, 且 $E(\boldsymbol{e}) = 0$, 所以 $E\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{L}$, 注意到 $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1$, 所以

$$\begin{aligned}E\hat{\boldsymbol{L}} &= E(\boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{12}) = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L} + \boldsymbol{A}_{12}\\ &= [(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} - (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R})^{-1}\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}]\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\\ &\quad + (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{R})^{-1} = \boldsymbol{L}.\end{aligned}$$

即 $\hat{\boldsymbol{L}} = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{12}$ 为 $\boldsymbol{L}$ 的无偏估计.

设 $\boldsymbol{A}\boldsymbol{Y}$ 为 $\boldsymbol{L}$ 的任一无偏估计, 即 $E(\boldsymbol{A}\boldsymbol{Y}) = \boldsymbol{L}$, 则有

$$\mathrm{cov}(\boldsymbol{A}\boldsymbol{Y}) = \sigma^2\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}, \quad \mathrm{cov}(\hat{\boldsymbol{L}}) = \sigma^2\boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A}_{11}^{\mathrm{T}}.$$

所以

$$\mathrm{cov}(\boldsymbol{A}\boldsymbol{Y}) - \mathrm{cov}(\hat{\boldsymbol{L}}) = \sigma^2(\boldsymbol{A} - \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}})(\boldsymbol{A} - \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}})^{\mathrm{T}} \geqslant 0.$$

从而 $\hat{\boldsymbol{L}}$ 为 $\boldsymbol{L}$ 的最小方差线性无偏估计. 证毕.

5.6.3 组合预测的权系数显著性检验

通过参数估计获得的组合预测的权系数后, 若发现有的预测方法的权系数接近零, 则必须对权系数进行显著性检验, 即检验假设

$$H : l_i = 0, \quad 1 \leqslant i \leqslant m.$$

若 H 被接受了, 则表明第 i 种预测方法与组合预测无显著关系, 此时可以去掉第 i 种预测方法.

因为 $\hat{\boldsymbol{L}} = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} + \boldsymbol{A}_{12}$, 所以 $\hat{\boldsymbol{L}}$ 服从多元正态分布 $N(\boldsymbol{L}, \sigma^2\boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A}_{11}^{\mathrm{T}})$.

记 $\boldsymbol{C} = (c_{ij})_{m\times m} = \boldsymbol{A}_{11}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A}_{11}^{\mathrm{T}}$, 则 $\hat{l}_i$ 服从一元正态分布 $N(l_i, \sigma^2 c_{ii})$, 令

$$\hat{\sigma}^2 = \|\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{L}}\|^2/(N - m - 1),$$

$\hat{\sigma}^2$ 为 σ^2 的一个估计. 下面构造 t 统计量. 令

$$T_i = \hat{l}_i/\hat{\sigma}\sqrt{c_{ii}}.$$

当假设 H 成立时, T_i 服从自由度为 $(N-m-1)$ 的 t 分布. 检验假设 H 的拒绝域为

$$W_\alpha = \left\{t \middle| |t| \geqslant t_{\frac{\alpha}{2}}(N - m - 1)\right\},$$

$\alpha \in (0,1)$ 为显著性水平, 因此, 当 $|T_i| < t_{\frac{\alpha}{2}}(N - m - 1)$ 时, 则接受假设 H, 可以去掉第 i 种预测方法.

第 6 章　非线性加权平均的最优组合预测的有效性理论

6.1　基于 L_2 和 L_1 范数的加权几何平均组合预测方法

根据组合预测模型的组合形式不同, 组合预测一般可分为线性组合预测和非线性组合预测. 加权算术平均组合预测属于线性组合预测, 而加权调和平均组合预测和加权几何平均组合预测属于非线性组合预测. 有时加权几何平均组合预测可以取得比加权算术平均组合预测更好的结果[80]. 目前提出的加权几何平均组合预测模型的参数估计均是基于预测误差向量的 L_2 范数最小为准则的. 本节建立基于对数误差向量 L_1 和 L_2 范数的加权几何平均组合预测模型.

6.1.1　基于 L_2 范数的加权几何平均的组合预测模型

加权几何平均组合预测是常用的组合预测方法之一. 它有时比加权算术平均组合预测方法有更好的组合预测效果. 下面给出其模型, 并讨论组合预测权系数的估计.

根据加权几何平均数计算公式, 令

$$\hat{x}_t = \prod_{i=1}^{m} x_{it}^{l_i}, \tag{6.1.1}$$

其中 $\hat{x}_t$ 为第 t 时刻的加权几何平均组合预测值, x_{it} 为第 i 种单项预测模型在第 t 时刻的预测值, $i=1,2,\cdots,m, t=1,2,\cdots,N, l_i$ 为第 i 种单项预测方法的加权系数, 且 $\sum_{i=1}^{m} l_i = 1, l_i \geqslant 0, i=1,2,\cdots,m.$

为了计算加权几何平均组合预测的权系数, 将式 (6.1.1) 两边取自然对数得

$$\ln\hat{x}_t = \sum_{i=1}^{m} l_i \ln x_{it}, \quad t=1,2,\cdots,N. \tag{6.1.2}$$

若预测误差不存在, 即 $\hat{x}_t = x_t$, 则有

$$\ln x_t = \sum_{i=1}^{m} l_i \ln x_{it}, \quad t=1,2,\cdots,N. \tag{6.1.3}$$

然而由于多种因素的影响, 组合预测的误差一般是不可避免的. 因此定义如下形式的预测误差.

定义 6.1.1 令

$$J=\sum_{t=1}^{N}e_t^2=\sum_{t=1}^{N}\left(\ln x_t-\sum_{i=1}^{m}l_i\ln x_{it}\right)^2, \tag{6.1.4}$$

则称 $J=\sum_{t=1}^{N}e_t^2$ 为基于 L_2 范数的加权几何平均的组合预测对数误差.

基于 L_2 范数的加权几何平均的组合预测对数误差越小表示预测越精确, 因此, 基于 L_2 范数的加权几何平均的组合预测模型为下列最优化问题

$$\min J=\sum_{t=1}^{N}e_t^2,$$
$$\text{s.t.}\begin{cases}\sum_{i=1}^{m}l_i(\ln x_t-\ln x_{it})=e_t, t=1,2,\cdots,N,\\ \sum_{i=1}^{m}l_i=1, l_i\geqslant 0, i=1,2,\cdots,m.\end{cases} \tag{6.1.5}$$

记 $\boldsymbol{R}=(1,1,\cdots,1)^{\mathrm{T}}$, $\boldsymbol{R}$ 为元素全为 1 的 m 维列向量, $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$, $e_{it}=\ln x_t-\ln x_{it}$, $\boldsymbol{E}=\left(\sum_{t=1}^{N}e_{it}e_{jt}\right)_{m\times m}$, 其中 $\boldsymbol{E}$ 称为加权几何平均组合预测的对数误差信息矩阵. 则有

$$\sum_{t=1}^{N}e_t^2=\sum_{t=1}^{N}\left(\sum_{i=1}^{m}l_i(\ln x_t-\ln x_{it})\right)^2=\sum_{t=1}^{N}\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_je_{it}e_{jt}=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}.$$

于是, 模型 (6.1.5) 化为矩阵形式

$$\min J=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L},$$
$$\text{s.t.}\begin{cases}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}=\boldsymbol{1},\\ \boldsymbol{L}\geqslant\boldsymbol{0}.\end{cases} \tag{6.1.6}$$

这是一个二次规划问题. 利用 Kuhn-Tucker 条件可将其转化为线性规划模型来求解最优组合预测权系数向量.

6.1.2 基于 L₁ 范数的加权几何平均的组合预测模型

预测误差平方和作为预测精度的指标, 存在一定的缺陷. 这就是预测误差再平方后就会使产生预测误差 “放大” 或 “缩小” 的效应, 即若预测绝对误差的绝对值大

于 1 时, 它平方后比其更大; 若预测相对误差的绝对值小于 1 时, 它平方后比其更小. 考虑到预测误差向量 L_2 范数的缺陷, 有必要引进预测误差向量 L_1 范数这个指标, 即用预测误差绝对值和来刻画预测精度. 它的稳健性比预测误差平方和要好, 特别是当数据中有异常极端值时, 模型的参数估计较稳健.

定义 6.1.2　令

$$F=\sum_{t=1}^{N}|e_t|=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}l_ie_{it}\right|,\quad F_i=\sum_{t=1}^{N}|e_{it}|,\quad i=1,2,\cdots,m. \tag{6.1.7}$$

则称 F 为加权几何平均的组合预测值和指标序列实际值之间的基于 L_1 范数的对数误差. 称 F_i 为第 i 种单项预测值和指标序列实际值之间的基于 L_1 范数的对数误差.

显然, F 为各种预测方法的加权系数向量的函数, 所以 F 可记为 $F(\boldsymbol{L})$.

对组合预测方法而言, 在理想的情况下, 若没有预测误差, 则 $F(\boldsymbol{L})=0$, 然而预测误差是不可避免的, $F(\boldsymbol{L})$ 越小表示加权几何平均的组合预测方法越接近指标序列实际值, 从而它就越精确有效. 因此, 基于 L_1 范数的最小对数误差为准则的加权几何平均组合预测模型可表示成如下模型 [81]

$$\min F(\boldsymbol{L})=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}l_ie_{it}\right|,$$
$$\text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{m}l_i=1,\\ l_i\geqslant 0,\quad i=1,2,\cdots,m.\end{cases} \tag{6.1.8}$$

从模型 (6.1.8) 的目标函数可知, 由于绝对值的存在使得模型 (6.1.8) 的目标函数不可导, 同时组合预测权系数 $l_1,l_2,\cdots,l_m$ 所在的位置也很分散, 因此可以通过如下变换, 去掉模型 (6.1.8) 的目标函数中的绝对值, 将其转化为一个线性规划问题进行求解. 令

$$\varepsilon_t=\sum_{i=1}^{m}l_ie_{it},\quad \varepsilon_t^+=\begin{cases}\varepsilon_t, & \text{当}\varepsilon_t\geqslant 0\\ 0, & \text{当}\varepsilon_t\leqslant 0\end{cases},\quad \varepsilon_t^-=\begin{cases}-\varepsilon_t, & \text{当}\varepsilon_t\leqslant 0\\ 0, & \text{当}\varepsilon_t\geqslant 0\end{cases},$$

则有

$$|\varepsilon_t|=\varepsilon_t^++\varepsilon_t^-,\quad \varepsilon_t=\varepsilon_t^+-\varepsilon_t^-,\quad \varepsilon_t^+\varepsilon_t^-=0.$$

在上述记号下, 模型 (6.1.8) 写成如下模型

$$
\min F(\boldsymbol{L})=\sum_{t=1}^{N}(\varepsilon_t^{+}+\varepsilon_t^{-}),
$$

$$
\text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{m} l_i e_{it}-\varepsilon_t^{+}+\varepsilon_t^{-}=0,\\ \displaystyle\sum_{i=1}^{m} l_i=1,\\ \varepsilon_t^{+}\geqslant 0,\varepsilon_t^{-}\geqslant 0,t=1,2,\cdots,N,\\ l_i\geqslant 0,i=1,2,\cdots,m.\end{cases} \tag{6.1.9}
$$

这是一个线性规划问题, 可用大 M 法求解, 也有现成的线性规划软件来求解. 模型 (6.1.9) 的最优解 $\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$ 为基于 L_1 范数的最小对数误差为准则的加权几何平均组合预测权系数.

6.1.3 实例分析

为了反映所提出的组合预测 (6.1.8) 的有效性, 按照预测效果评价原则, 选择 4.7.1 节组合预测效果评价的指标体系.

应用文献 [82] 的数据进行实例分析. 某省 11 个年份社会商品零售总额的统计资料及采用指数曲线预测方法和抛物线预测方法这两种单项预测模型的预测数值见表 6.1.1.

表 6.1.1 社会商品零售总额及两种单项预测方法预测值

年份t	零售总额 (亿元)	指数曲线预测方法预测值x_{1t}	抛物线预测方法预测值x_{2t}
1	57.0	54.52	64.48
2	65.4	62.89	66.74
3	75.4	72.54	68.72
4	82.5	83.67	76.61
5	92.8	96.51	88.42
6	102.7	111.32	104.15
7	119.5	128.41	123.79
8	143.8	148.11	147.35
9	169.7	170.84	174.82
10	201.0	197.06	206.21
11	215.2	227.31	241.51

由定义 6.1.2 计算得两种单项预测方法各个时刻预测值与相应的实际值之间对数误差, 见表 6.1.2.

表 6.1.2　两种单项预测方法对数误差及加权几何平均组合预测值

年份t	指数曲线预测方法对数误差e_{1t}	抛物线预测法对数误差e_{2t}	L_1范数的最小对数误差准则的加权几何平均组合预测值	加权几何平均组合预测绝对误差
1	0.0445	−0.1233	57.00	0.0
2	0.0391	−0.0203	63.88	1.51
3	0.0387	0.0928	71.50	3.89
4	−0.0141	0.0741	81.73	0.76
5	−0.0392	0.0483	94.29	−1.49
6	−0.0806	−0.0140	109.37	−6.67
7	−0.0719	−0.0353	127.16	−7.66
8	−0.0295	−0.0244	147.91	−4.10
9	−0.0067	−0.0297	171.88	−2.18
10	0.0198	−0.0256	199.44	1.55
11	−0.0547	−0.1153	230.99	−15.79

所以将这些数据代入模型 (6.1.8) 中可得如下最优化模型

$$\begin{aligned}\min F(l_1,l_2) = &|0.0445l_1 - 0.1233l_2| + |0.0391l_1 - 0.0203l_2| + |0.0387l_1 + 0.0928l_2| \\ &+ |-0.0141l_1 + 0.0741l_2| + |-0.0392l_1 + 0.0483l_2| + |-0.0806l_1 \\ &-0.0140l_2| + |-0.0719l_1 - 0.0353l_2| + |-0.0295l_1 - 0.0244l_2| \\ &+ |-0.0067l_1 - 0.0297l_2| + |0.0198l_1 - 0.0256l_2| + |-0.0547l_1 \\ &-0.1153l_2|,\end{aligned}$$

$$\text{s.t.}\begin{cases} l_1 + l_2 = 1, \\ l_1 \geqslant 0, l_2 \geqslant 0.\end{cases}$$

利用 Matlab 最优化工具箱计算得上述最优化模型的解为

$$l_1^* = 0.7348, \quad l_2^* = 0.2652.$$

将模型 (6.1.8) 的最优解代入到 (6.1.7) 式中, 计算得 L_1 范数的最小对数误差准则的加权几何平均组合预测值及其绝对预测误差, 见表 6.1.2.

按照预测效果的评价指标体系, 计算组合预测模型 (6.1.7) 的几个指标得 (1) 平方和误差, SSE= 397.0450; (2) 均方误差, MSE=1.8115; (3) 平均绝对误差, MAE= 4.1495; (4) 平均绝对百分比误差, MAPE=0.0029; (5) 均方百分比误差, MSPE= 0.0011. 文献 [82] 中给出的是基于误差平方和的加权几何平均组合预测模型, 其计算最优解为

$$l_1^{**} = 0.5687, \quad l_2^{**} = 0.4313.$$

文献 [82] 只给出两个预测效果的评价指标. 即平方和误差 SSE=449.41, 均方百分比

误差 MSPE=0.0122, 可见就这两个评价指标而言, 本节提出的模型的结果优于文献 [82] 的计算结果.

另外, 组合预测模型 (6.1.8) 的 L_1 范数最小对数误差 $F(l_1^*,l_2^*)=\sum_{t=1}^{11}\left|\sum_{i=1}^{2}l_i^*e_{it}\right|=0.3463$, 而指数曲线预测方法的 L_1 范数的对数误差 $F_1=\sum_{t=1}^{11}|e_{1t}|=0.4388$, 抛物线预测方法的 L_1 范数的对数误差 $F_2=\sum_{t=1}^{11}|e_{2t}|=0.6031$, 所以 $F(l_1^*,l_2^*)<\min\{F_1,F_2\}$, 组合预测模型 (6.1.8) 为优性组合预测. 从而表明提出的组合预测方法的有效性.

6.2 基于 L1 范数的加权几何平均组合预测方法的性质

6.2.1 几个概念 [83]

本节针对该模型提出新的优性组合预测、预测方法优超、冗余度等概念, 研究它的性质, 并得出一些有益的结论. 从而在理论上说明基于 L_1 范数的加权几何平均的组合预测方法的有效性.

记 $F_{\min}=\min\{F_i,i=1,2,\cdots,m\}$, $F_{\max}=\max\{F_i,i=1,2,\cdots,m\}$, 即 $F_{\min}$, $F_{\max}$ 分别表示 m 种单项预测值和指标实际值序列之间最小的和最大的基于 L_1 范数的对数误差, F_i 的意义参见定义 6.1.2 中的式 (6.1.7).

定义 6.2.1 若 $F(\boldsymbol{L})<F_{\min}$, 则称组合预测模型 (6.1.8) 为优性组合预测, 若 $F_{\min}\leqslant F(\boldsymbol{L})\leqslant F_{\max}$, 则称之为非劣性组合预测, 若 $F(\boldsymbol{L})>F_{\max}$, 则称之为劣性组合预测.

定义 6.2.1 表明, 只有加权几何平均的组合预测的基于 L_1 范数的对数误差小于各单项预测方法中的最小者, 则该组合预测模型才为优性组合预测. 即从 L_1 范数的最小对数误差这个角度考虑, 优性组合预测要 "好" 于各单项预测方法中的最 "好" 者.

定义 6.2.2 若第 j 种和第 k 种单项预测方法预测对数误差满足不等式

$$|e_{jt}|\leqslant|e_{kt}|,t=1,2,\cdots,N,\text{且对某个时刻}t_0\text{成立}\,|e_{jt_0}|<|e_{kt_0}|,t_0\in\{1,2,\cdots,N\}.$$

则称第 j 种单项预测方法优超第 k 种单项预测方法. 若对任意时刻均有严格的不等号成立, 则称第 j 种单项预测方法严格优超第 k 种单项预测方法.

实际上, 根据定义 6.2.2 知, 若第 j 种单项预测方法优超第 k 种单项预测方法, 则有 $F_j<F_k$. 即就 L_1 范数的最小对数误差这个准则而言, 第 j 种单项预测方法比第 k 种更接近指标序列实际值, 因此可认为第 j 种单项预测方法要优于第 k 种单项预测方法.

定义 6.2.3 若某种单项预测方法在组合预测模型 (6.1.8) 中的最优权系数为零, 即它增加到加权几何平均的组合预测模型中不能减小其和指标序列实际值之间 L_1 范数的对数误差, 则称该种单项预测方法为冗余预测方法, 表明该种单项预测方法只提供冗余信息.

6.2.2 非劣性和优性组合预测存在性[83]

引理 6.2.1 组合预测和各单项预测方法与指标序列实际值之间的 L_1 范数的对数误差满足如下不等式

$$F(\boldsymbol{L}) \leqslant \sum_{i=1}^{m} l_i F_i. \tag{6.2.1}$$

证明 因为加权几何平均组合预测权系数向量 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 满足非负性, 所以由式 (6.1.7) 知

$$F(\boldsymbol{L}) = \sum_{t=1}^{N} \left| \sum_{i=1}^{m} l_i e_{it} \right| \leqslant \sum_{t=1}^{N} \sum_{i=1}^{m} l_i |e_{it}| = \sum_{i=1}^{m} l_i \left(\sum_{t=1}^{N} |e_{it}| \right) = \sum_{i=1}^{m} l_i F_i.$$

引理 6.2.1 表明, 组合预测方法与指标序列实际值之间 L_1 范数的对数误差小于或等于各单项预测方法的加权平均值, 且加权系数为组合预测权系数向量.

定理 6.2.1 组合预测模型 (6.1.8) 的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 设 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为组合预测模型 (6.1.8) 的任一个可行解, 则有

$$\sum_{i=1}^{m} l_i = 1, \quad l_i \geqslant 0, \quad i = 1, 2, \cdots, m.$$

由式 (6.2.1) 知

$$F(\boldsymbol{L}) \leqslant \sum_{i=1}^{m} l_i F_i \leqslant \sum_{i=1}^{m} l_i F_{\max} = F_{\max} \sum_{i=1}^{m} l_i = F_{\max}.$$

再由定义 6.2.1 得结论成立.

定理 6.2.1 表明, 就最小的 L_1 范数的对数误差这个准则而言, 组合预测模型 (6.1.8) 的任一个归一化非负权系数所对应的组合预测至少不会比 "最差" 的单项预测方法还要 "差".

推论 6.2.1 简单等权平均组合预测方法至少是非劣性组合预测.

定理 6.2.2 若 $\mathrm{rank}(\boldsymbol{E}) = 1, \boldsymbol{E}$ 的第 1 列存在非零元素, 且满足 $\sum_{i=1}^{m} e_{i1} = 0$, 则简单等权平均组合预测方法为最优性组合预测方法. 即 $l_1 = l_2 = \cdots = l_m = \dfrac{1}{m}$

为组合预测模型 (6.1.8) 的最优解. 其中 rank($\boldsymbol{E}$) 表示组合预测模型 (6.1.8) 的对数误差信息矩阵 $\boldsymbol{E}$ 的秩.

证明 对 $\boldsymbol{E}$ 按列进行分块, 设 $\boldsymbol{E}=(\boldsymbol{E}_1,\boldsymbol{E}_2,\cdots,\boldsymbol{E}_N)$, 其中 $\boldsymbol{E}_t=(e_{1t},e_{2t},\cdots,e_{mt})^{\mathrm{T}}$ 为 $\boldsymbol{E}$ 的第 t 列, $t=1,2,\cdots,N$. 因为 rank($\boldsymbol{E}$) $=1$, 所以存在比例常数 c_t, 使得 $\boldsymbol{E}_t=c_t\boldsymbol{E}_1\quad t=2,3,\cdots,N$, 即

$$e_{it}=c_te_{i1},\quad i=1,2,\cdots,m,\quad t=2,3,\cdots,N.$$

又因为 $\sum\limits_{i=1}^{m}e_{i1}=0$, 所以 $\sum\limits_{i=1}^{m}e_{it}=c_t\sum\limits_{i=1}^{m}e_{i1}=0,\ t=2,3,\cdots,N.$

当简单等权平均组合预测方法的权系数 $l_1=l_2=\cdots=l_m=\dfrac{1}{m}$ 时, 则有

$$F\left(\frac{1}{m},\frac{1}{m},\cdots,\frac{1}{m}\right)=\frac{1}{m}\sum_{t=1}^{N}\left|\sum_{i=1}^{m}e_{it}\right|=0.$$

即 $F\left(\dfrac{1}{m},\dfrac{1}{m},\cdots,\dfrac{1}{m}\right)$ 达到其下界 0. 所以简单等权平均组合预测方法为最优性组合预测方法.

定理 6.1.2 表明, 在一定的条件下简单等权平均组合预测方法也可能成为最优性组合预测方法. 下面给出基于 L₁ 范数的加权几何平均的组合预测是优性组合预测的一个充分条件.

定理 6.2.3 设 $F_1=\min\{F_i,i=1,2,\cdots,m\},b_t\in(-|e_{1t}|,|e_{1t}|),t=1,2,\cdots,N$, 若 $(N+1)\times m$ 的线性方程组 $\boldsymbol{AL}=\boldsymbol{\beta}$ 存在非负解, 则组合预测模型 (6.1.8) 的一定存在优性组合预测. 其中 $\boldsymbol{A}=(\boldsymbol{E},\boldsymbol{R})^{\mathrm{T}},\boldsymbol{E}=(e_{it})_{m\times N}$ 为组合预测模型 (6.1.8) 的对数误差信息矩阵, $\boldsymbol{R}=(1,1,\cdots,1)^{\mathrm{T}}$ 为元素全为 1 的 m 维列向量, $\boldsymbol{\beta}=(b_1,b_2,\cdots,b_N,1)^{\mathrm{T}}$ 为 $N+1$ 维列向量.

证明 设 $(N+1)\times m$ 的线性方程组 $\boldsymbol{AL}=\boldsymbol{\beta}$ 存在的非负解为 $\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$, 则有 $\boldsymbol{AL}^*=\boldsymbol{\beta}$, 利用矩阵乘法将其分量形式得

$$\sum_{i=1}^{m}e_{it}l_i^*=b_t,\quad t=1,2,\cdots,N,\quad \sum_{i=1}^{m}l_i^*=1.$$

因为 $b_t\in(-|e_{1t}|,|e_{1t}|)$, 所以有

$$\left|\sum_{i=1}^{m}l_i^*e_{it}\right|=|b_t|<|e_{1t}|,\quad t=1,2,\cdots,N. \tag{6.2.2}$$

由式 (6.1.7)、(6.2.2) 知

$$F(\boldsymbol{L}^*)=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}l_i^*e_{it}\right|<\sum_{t=1}^{N}|e_{1t}|=F_1.$$

又由条件知 $F_1 = \min\{F_i, i = 1, 2, \cdots, m\}$, 所以 $F(\boldsymbol{L}^*) < \min\{F_i, i = 1, 2, \cdots, m\}$. 由定义 6.2.1 知结论成立. 证毕.

推论 6.2.2 设 $F_1 = \min\{F_i, i = 1, 2, \cdots, m\}$, 若组合预测模型的对数误差信息矩阵 $\boldsymbol{E} = (e_{it})_{m\times N}$ 中任一列 m 个元素 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的算术平均数 $\bar{e}_t = \dfrac{1}{m}\sum\limits_{i=1}^{m} e_{it} \in (-|e_{1t}|, |e_{1t}|)$, $t = 1, 2, \cdots, N$, 则简单等权平均组合预测方法是优性组合预测方法.

证明 令 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}} = (1/m, 1/m, \cdots, 1/m)^{\mathrm{T}}$, 由条件 $\bar{e}_t = \dfrac{1}{m}\sum\limits_{i=1}^{m} e_{it} \in (-|e_{1t}|, |e_{1t}|)$ 知, 则它是满足线性方程组 $\boldsymbol{A}\boldsymbol{L}^* = \boldsymbol{\beta}$ 的非负解, 从而由定理 6.2.3 得结论成立. 证毕.

6.2.3 预测冗余信息的存在性及判定[83]

对于组合预测模型 (6.1.8) 有如下结论成立.

定理 6.2.4 组合预测模型 (6.1.8) 的最优目标函数值是 m 的单调不增函数, 即

$$F^*(l_1, l_2, \cdots, l_m) \geqslant F^*(l_1, l_2, \cdots, l_{m+1}), \tag{6.2.3}$$

其中 m 为参与组合预测的各单项预测方法总个数, $F^*(l_1, l_2, \cdots, l_m)$ 表示 m 个单项预测方法参与的组合预测模型 (6.1.8) 对应的最小的基于 L_1 范数的对数误差. $F^*(l_1, l_2, \cdots, l_{m+1})$ 表示再增加一个单项预测方法的组合预测模型对应的最小的基于 L_1 范数的对数误差.

证明 设 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$ 为 m 个单项预测方法参与的组合预测模型 (6.1.8) 的最优解, 则

$$F^*(l_1, l_2, \cdots, l_m) = \sum_{t=1}^{N}\left|\sum_{i=1}^{m} l_i^* e_{it}\right|,$$

其中 $\sum\limits_{i=1}^{m} l_i^* = 1, l_i^* \geqslant 0, i = 1, 2, \cdots, m+1$.

设 $\bar{\boldsymbol{L}} = (\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1})^{\mathrm{T}}$ 为再增加一个单项预测方法, 共 $m+1$ 个单项预测方法参与的组合预测模型 (6.1.8) 的最优解, 则有

$$F^*(l_1, l_2, \cdots, l_m, l_{m+1}) = \sum_{t=1}^{N}\left|\sum_{i=1}^{m+1} \bar{l}_i e_{it}\right|,$$

其中 $\sum\limits_{i=1}^{m+1} \bar{l}_i = 1$, $\bar{l}_i \geqslant 0$, $i = 1, 2, \cdots, m+1$.

令 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*,0)^{\mathrm{T}}$, 显然, $\boldsymbol{L}$ 为 $m+1$ 个单项预测方法参与的组合预测模型 (6.1.8) 的可行解, 则有

$$F^*(l_1,l_2,\cdots,l_m,l_{m+1})\leqslant F(l_1^*,l_2^*,\cdots,l_m^*,0).$$

而

$$F(l_1^*,l_2^*,\cdots,l_m^*,0)=\sum_{t=1}^N\left|\sum_{i=1}^m l_i^* e_{it}+0\times e_{(m+1)t}\right|=F^*(l_1,l_2,\cdots,l_m).$$

所以 $F^*(l_1,l_2,\cdots,l_{m+1})\leqslant F^*(l_1,l_2,\cdots,l_m)$. 证毕.

定理 6.2.4 指出, 当再增加一个单项预测方法时, 组合预测模型 (6.1.8) 对应的最小的 L_1 范数的对数误差可能不变. 这表明组合预测模型 (6.1.8) 可能存在冗余预测方法. 下面一个定理为冗余信息提供了判定.

定理 6.2.5 若对数误差信息矩阵 $\boldsymbol{E}$ 的任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号相同, $t=1,2,\cdots,N$, 且第 j 种单项预测方法优超第 k 种单项预测方法, 则组合预测模型 (6.1.8) 的冗余度至少为 $1/m$. 即至少存在第 k 种单项预测方法为冗余预测方法.

证明 采用反证法. 假设第 k 种单项预测方法不为冗余预测方法, 即在组合预测模型 (6.1.8) 的最优解 $\boldsymbol{L}^*=(l_1^*,\cdots,l_j^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 中, 第 k 种单项预测方法对应的最优权系数 $l_k^*>0$.

因为任意第 t 列元素 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t=1,2,\cdots,N$, 并注意到最优解 $\boldsymbol{L}^*$ 满足非负性, 所以 $\left|\sum_{i=1}^m l_i^* e_{it}\right|=\sum_{i=1}^m l_i^*\left|e_{it}\right|, t=1,2,\cdots,N$, 从而最优解 $\boldsymbol{L}^*$ 对应的目标函数值为

$$F(\boldsymbol{L}^*)=\sum_{t=1}^N\left|\sum_{i=1}^m l_i^* e_{it}\right|=\sum_{t=1}^N\sum_{i=1}^m l_i^*\left|e_{it}\right|=\sum_{i=1}^m l_i^* F_i. \tag{6.2.4}$$

令 $\hat{\boldsymbol{L}}=(l_1^*,\cdots,l_j^*+l_k^*,\cdots,0,\cdots,l_m^*)^{\mathrm{T}}$, 即 $\hat{\boldsymbol{L}}$ 的第 j 个分量为 $l_j^*+l_k^*$, 第 k 个分量为 0, 其他 $m-2$ 个分量与 $\boldsymbol{L}^*$ 相同, 显然 $\hat{\boldsymbol{L}}$ 为组合预测模型 (6.1.8) 的一个可行解, 同理, 对应的目标函数值为

$$\begin{aligned}F(\hat{\boldsymbol{L}})&=\sum_{t=1}^N\left(\sum_{i\neq j,k} l_i^*\left|e_{it}\right|+(l_j^*+l_k^*)\left|e_{jt}\right|+0\times\left|e_{kt}\right|\right)\\&=\sum_{t=1}^N\left(\sum_{i=1}^m l_i^*\left|e_{it}\right|+l_k^*(\left|e_{jt}\right|-\left|e_{kt}\right|)\right)=\sum_{i=1}^m l_i^* F_i+l_k^*(F_j-F_k).\end{aligned} \tag{6.2.5}$$

因为第 j 种单项预测方法优超第 k 种单项预测方法, 由定义 6.2.2 知

$$\left|e_{jt}\right|\leqslant\left|e_{kt}\right|,\quad t=1,2,\cdots,N,\quad \left|e_{jt_0}\right|<\left|e_{kt_0}\right|,\quad t_0\in\{1,2,\cdots,N\}, \tag{6.2.6}$$

由式 (6.1.7)、(6.2.5) 得

$$F_j < F_k. \tag{6.2.7}$$

由式 (6.2.4)、(6.2.5)、(6.2.7) 并注意到假设 $l_k^* > 0$, 所以有

$$F(\hat{\boldsymbol{L}}) < F(\boldsymbol{L}^*).$$

而这与 $\boldsymbol{L} = (l_1^*, \cdots, l_j^*, \cdots, l_k^*, \cdots, l_m^*)^{\mathrm{T}}$ 为组合预测模型 (6.1.8) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 即组合预测模型 (6.1.8) 的冗余度至少为 $1/m$. 证毕.

6.3 调和平均的组合预测方法的性质

6.3.1 基于误差平方和准则的调和平均组合预测模型

调和平均组合预测是一种比较常用的组合预测方法, 实证研究表明, 在某些情况下它可以比加权算术平均取得更好的组合预测效果. 因此有必要对其进行研究, 以便选择适当的组合预测模型.

设某一预测问题的指标序列为 $\{x_t, t = 1, 2, \cdots, N\}$, 对此预测问题存在 m 种单项预测方法对其进行预测, 设 x_{it} 为第 i 种单项预测模型在第 t 时刻的预测值, $i = 1, 2, \cdots, m, t = 1, 2, \cdots, N$. 根据加权调和平均数计算公式, 令

$$\hat{x}_t = \frac{\sum_{i=1}^{m} l_i}{\sum_{i=1}^{m} \frac{l_i}{x_{it}}} = \frac{1}{\sum_{i=1}^{m} \frac{l_i}{x_{it}}}, \tag{6.3.1}$$

其中 $\hat{x}_t$ 第 t 时刻的加权调和平均组合预测值, l_i 为第 i 种单项预测方法的加权系数, 且 l_i 满足 $\sum_{i=1}^{m} l_i = 1$, $l_i \geqslant 0, i = 1, 2, \cdots, m$.

为了计算加权调和平均组合预测的权系数可将式 (6.3.1) 变化为

$$\sum_{i=1}^{m} \frac{l_i}{x_{it}} = \frac{1}{\hat{x}_t}, \quad t = 1, 2, \cdots, N. \tag{6.3.2}$$

假设在理想的情况下, $\hat{x}_t = x_t$, 即预测误差的不存在, 则有

$$\sum_{i=1}^{m} \frac{l_i}{x_{it}} = \frac{1}{x_t}, \quad t = 1, 2, \cdots, N. \tag{6.3.3}$$

然而在预测的实践中, 组合预测的误差一般是存在的. 因此定义如下形式的预测误差

$$e_t = \sum_{i=1}^{m} \frac{l_i}{x_{it}} - \frac{1}{x_t}, \quad t = 1, 2, \cdots, N, \tag{6.3.4}$$

其中 e_t 称为加权调和平均组合预测的在第 t 时刻的组合预测误差.

显然, 组合预测的预测误差平方和越小越好, 因此, 以预测误差平方和为准则的非负权系数的组合预测模型为下列最优化问题

$$\min J = \sum_{t=1}^{N} e_t^2,$$
$$\text{s.t.} \begin{cases} \displaystyle\sum_{i=1}^{m} \frac{l_i}{x_{it}} - \frac{1}{x_t} = e_t, t = 1, 2, \cdots, N, \\ \displaystyle\sum_{i=1}^{m} l_i = 1, l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{6.3.5}$$

记 $\boldsymbol{R} = (1, 1, \cdots, 1)^{\mathrm{T}}$, $\boldsymbol{R}$ 为元素全为 1 的 m 维列向量, $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$, $\boldsymbol{E} = (e_1, e_2, \cdots, e_N)^{\mathrm{T}}$,

$$\boldsymbol{X} = (1/x_1, 1/x_2, \cdots, 1/x_N)^{\mathrm{T}}, \quad \boldsymbol{F} = \begin{pmatrix} 1/x_{11} & 1/x_{21} & \cdots & 1/x_{m1} \\ 1/x_{12} & 1/x_{22} & \cdots & 1/x_{m2} \\ \vdots & \vdots & & \vdots \\ 1/x_{1N} & 1/x_{2N} & \cdots & 1/x_{mN} \end{pmatrix}^{\mathrm{T}}.$$

则有

$$\sum_{t=1}^{N} e_t^2 = \boldsymbol{E}^{\mathrm{T}}\boldsymbol{E} = (\boldsymbol{FL} - \boldsymbol{X})^{\mathrm{T}}(\boldsymbol{FL} - \boldsymbol{X}).$$

于是模型 (6.3.5) 化为矩阵形式

$$\min J = (\boldsymbol{FL} - \boldsymbol{X})^{\mathrm{T}}(\boldsymbol{FL} - \boldsymbol{X}),$$
$$\text{s.t.} \begin{cases} \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L} = 1, \\ \boldsymbol{L} \geqslant \boldsymbol{0}. \end{cases} \tag{6.3.6}$$

显然, 这是一个二次规划问题. 根据二次规划理论可知, 该二次规划问题的最优解一定存在. 利用 Kuhn-Tucker 条件可将其转化为线性规划模型来求解, 从而可获得最优组合预测权系数向量.

6.3.2 基于几何距离准则的调和平均组合预测模型几个概念[84]

令 $\tilde{x}_t = 1/x_t, \tilde{x}_{it} = 1/x_{it}, t = 1, 2, \cdots, N, \tilde{x}_t$ 和 $\tilde{x}_{it}$ 分别为实际观察值和预测值的倒数, 则 (6.3.3) 式可写成

$$\sum_{i=1}^{m} l_i \tilde{x}_{it} - \tilde{x}_t = 0, \quad t = 1, 2, \cdots, N. \tag{6.3.7}$$

从解析几何角度考虑, 式 (6.3.7) 代表了 N 个以 $l_1, l_2, \cdots, l_m$ 为坐标空间的平面方程. 由于组合预测拟合误差是不可避免的, 一般不存在一组组合预测权系数同时满足式 (6.3.7) 的 N 个平面方程. 因此确定组合预测权系数的一个自然的想法是, 在以 $l_1, l_2, \cdots, l_m$ 为坐标空间中找到某点, 使得该点到 N 个平面的几何距离总和为最短[85]. 为此引进如下概念.

定义 6.3.1 令

$$d = \sum_{t=1}^{N} \left| \left(\sum_{i=1}^{m} l_i \tilde{x}_{it} - \tilde{x}_t \right) \Bigg/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2} \right|, \tag{6.3.8}$$

$$d_i = \sum_{t=1}^{N} \left| (\tilde{x}_{it} - \tilde{x}_t) \Bigg/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2} \right|, \quad i = 1, 2, \cdots, m. \tag{6.3.9}$$

则称 d 为加权调和平均的组合预测值和实际观察值之间的基于 L_1 范数的几何距离. 称 d_i 为第 i 种单项预测值和指标序列实际观察值之间的基于 L_1 范数的几何距离.

根据点到平面的距离公式知, d 实际上为组合预测的加权向量 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 到式 (6.3.7) 的 N 个平面方程的几何距离之和. 显然 d 为各种预测方法的加权系数向量的函数, 所以 d 可记为 $d(\boldsymbol{L})$. 特别地, 当加权系数向量为第 i 个 m 维标准单位列向量时, 即 $\boldsymbol{L}_i = (0, \cdots, 0, 1, 0, \cdots, 0)^{\mathrm{T}}$ 的第 i 个分量为 1, 其余 $m-1$ 个分量为 0, 此时组合预测方法就转变为第 i 种单项预测方法, 从而 $d_i = d(\boldsymbol{L}_i)$ 为第 i 种单项预测值和指标序列实际观察值之间的距离, $i = 1, 2, \cdots, m$. 为方便起见, 引进如下定义.

定义 6.3.2 令 $e_{it} = \tilde{x}_{it} - \tilde{x}_t$, 则称 e_{it} 为第 i 种预测方法第 t 时刻预测值与相应的实际观察值之间倒数误差, $i = 1, 2, \cdots, m, t = 1, 2, \cdots, N$, 称矩阵 $\boldsymbol{E} = (e_{it})_{m \times N}$ 为加权调和平均的组合预测模型的倒数误差信息矩阵.

因为加权调和平均组合预测权系数向量 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 满足归一性和非负性, 所以将式 (6.3.8)、(6.3.9) 变为

$$d(\boldsymbol{L}) = \sum_{t=1}^{N} \left| \left(\sum_{i=1}^{m} l_i (\tilde{x}_{it} - \tilde{x}_t) \right) \Bigg/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2} \right| = \sum_{t=1}^{N} \left| \sum_{i=1}^{m} l_i e_{it} \Bigg/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2} \right|, \tag{6.3.10}$$

$$d_i = \sum_{t=1}^{N} \left| e_{it} \Big/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2} \right|, \quad i = 1, 2, \cdots, m, \tag{6.3.11}$$

对组合预测方法而言, $d(\boldsymbol{L})$ 越小表示加权调和平均的组合预测方法越接近实际观察值, 从而它就越精确有效. 因此, 基于 L_1 范数的最小几何距离为准则的加权调和平均组合预测模型可表示为[85]

$$\min d(\boldsymbol{L}) = \sum_{t=1}^{N} \left| \sum_{i=1}^{m} l_i e_{it} \Big/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2} \right|,$$
$$\text{s.t.} \begin{cases} \sum_{i=1}^{m} l_i = 1, \\ l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{6.3.12}$$

记 $d_{\min} = \min\{d_i, i = 1, 2, \cdots, m\}, d_{\max} = \max\{d_i, i = 1, 2, \cdots, m\}$.

定义 6.3.3 若 $d(\boldsymbol{L}) < d_{\min}$, 则称组合预测为优性组合预测, 若 $d_{\min} \leqslant d(\boldsymbol{L}) \leqslant d_{\max}$, 则称它为非劣性组合预测, 若 $d(\boldsymbol{L}) > d_{\max}$, 则称它为劣性组合预测.

定义 6.3.3 表明, 从最小几何距离考虑, 优性组合预测要 "好" 于各单项预测方法中的最 "好" 者.

定义 6.3.4 若第 j 种和第 k 种单项预测方法倒数误差满足 $|e_{jt}| \leqslant |e_{kt}|, t = 1, 2, \cdots, N$, 且对某个时刻 t_0 成立 $|e_{jt_0}| < |e_{kt_0}|, t_0 \in \{1, 2, \cdots, N\}$, 则称第 j 种单项预测方法优超第 k 种. 若对任意时刻均有严格的不等号成立, 则称第 j 种单项预测方法严格优超第 k 种.

实际上, 根据式 (6.3.11) 知, 第 j 种单项预测方法优超第 k 种已蕴涵 $d_j < d_k$. 即就最小几何距离而言, 第 j 种单项预测方法比第 k 种更接近实际观察值, 因此可认为第 j 种要优于第 k 种.

定义 6.3.5 若某种单项预测方法增加到组合预测中不能减小组合预测和实际观察值之间的几何距离, 则称该单项预测方法为冗余预测方法. 即该种单项预测方法在组合预测模型的最优权系数为零, 表明该预测方法只提供冗余信息.

6.3.3 非劣性组合预测和优性组合预测存在的条件

引理 6.3.1 组合预测和各单项预测方法的预测值与指标序列实际观察值之间的几何距离满足如下不等式

$$d(\boldsymbol{L}) \leqslant \sum_{i=1}^{m} l_i d_i. \tag{6.3.13}$$

证明 因为加权调和平均组合预测权系数向量满足非负性, 所以由式 (6.3.10)、(6.3.11) 知

$$
\begin{aligned}
d(\boldsymbol{L}) &= \sum_{t=1}^{N}\left|\sum_{i=1}^{m} l_i e_{it} \Big/ \sqrt{\sum_{i=1}^{m}\tilde{x}_{it}^2}\right| \leqslant \sum_{t=1}^{N}\sum_{i=1}^{m} l_i \left|e_{it} \Big/ \sqrt{\sum_{i=1}^{m}\tilde{x}_{it}^2}\right| \\
&= \sum_{i=1}^{m} l_i \left(\sum_{t=1}^{N}\left|e_{it} \Big/ \sqrt{\sum_{i=1}^{m}\tilde{x}_{it}^2}\right|\right) = \sum_{i=1}^{m} l_i d_i.
\end{aligned}
$$

证毕.

该引理 6.3.1 表明, 组合预测值与实际观察值之间的几何距离小于或等于各单项预测方法与实际观察值之间的几何距离的加权平均, 且权系数为组合预测权系数向量.

定理 6.3.1 组合预测 (6.3.12) 式的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 设 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为 (6.3.12) 式的任一个可行解, 则 $\boldsymbol{L}$ 满足归一化约束条件. 由式 (6.3.13) 知

$$
d(\boldsymbol{L}) \leqslant \sum_{i=1}^{m} l_i d_i \leqslant \sum_{i=1}^{m} l_i d_{\max} = d_{\max}\sum_{i=1}^{m} l_i = d_{\max}.
$$

再由定义 6.3.3 得结论成立. 证毕.

定理 6.3.1 表明, 就最小几何距离角度, 组合预测式 (6.3.12) 的任一个归一化非负权系数所对应的组合预测至少不会比 “最差” 的单项预测方法还要 “差”. 显然有推论 6.3.1 成立.

推论 6.3.1 简单等权平均组合预测方法至少是非劣性组合预测.

定理 6.3.2 若组合预测的倒数误差信息矩阵 $\boldsymbol{E} = (e_{it})_{m\times N}$ 的秩为 1, $\boldsymbol{E}$ 的第 1 列存在非零元素, 且满足 $\sum_{i=1}^{m} e_{i1} = 0$, 则简单等权平均组合预测方法为最优性组合预测方法. 即 $l_1 = l_2 = \cdots = l_m = 1/m$ 为组合预测式 (6.3.12) 的最优解.

证明 对 $\boldsymbol{E}$ 按列进行分块, 设 $\boldsymbol{E} = (\boldsymbol{E}_1, \boldsymbol{E}_2, \cdots, \boldsymbol{E}_N)$, 其中 $\boldsymbol{E}_t = (e_{1t}, e_{2t}, \cdots, e_{mt})^{\mathrm{T}}$ 为 $\boldsymbol{E}$ 的第 t 列, $t = 1, 2, \cdots, N$, 因为 $\boldsymbol{E} = (e_{it})_{m\times N}$ 的秩为 1, 所以存在比例常数 k_t, 使得 $\boldsymbol{E}_t = k_t \boldsymbol{E}_1 \quad t = 2, 3, \cdots, N$, 即

$$
e_{it} = k_t e_{i1}, \quad i = 1, 2, \cdots, m, \quad t = 2, 3, \cdots, N.
$$

又因为 $\sum_{i=1}^{m} e_{i1} = 0$, 所以 $\sum_{i=1}^{m} e_{it} = k_t \sum_{i=1}^{m} e_{i1} = 0, t = 2, 3, \cdots, N.$

当简单等权平均组合预测方法的权系数 $l_1 = l_2 = \cdots = l_m = 1/m$ 时, 则有

$$
d(1/m, 1/m, \cdots, 1/m) = \frac{1}{m}\sum_{t=1}^{N}\left|\sum_{i=1}^{m} e_{it} \Big/ \sqrt{\sum_{i=1}^{m}\tilde{x}_{it}^2}\right| = 0.
$$

即 $d(1/m, 1/m, \cdots, 1/m)$ 达到其下界 0. 所以简单等权平均组合预测方法为最优性组合预测方法. 证毕.

下面给出组合预测式 (6.3.12) 的优性组合预测存在的一个充分条件.

定理 6.3.3 设 $d_1 = \min\{d_i, i = 1, 2, \cdots, m\}, b_t \in (-|e_{1t}|, |e_{1t}|), t = 1, 2, \cdots, N$, 若 $(N+1) \times m$ 的线性方程组 $\boldsymbol{AL} = \boldsymbol{b}$ 存在非负解, 则组合预测式 (6.3.12) 一定存在优性组合预测. 其中 $\boldsymbol{A} = (\boldsymbol{E}, \boldsymbol{R})^{\mathrm{T}}$, $\boldsymbol{E} = (e_{it})_{m\times N}$ 为倒数误差信息矩阵, $\boldsymbol{R} = (1, 1, \cdots, 1)^{\mathrm{T}}$ 为元素全为 1 的 m 维列向量, $\boldsymbol{b} = (b_1, b_2, \cdots, b_N, 1)^{\mathrm{T}}$ 为 $N+1$ 维列向量, $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为组合预测的加权向量.

证明 设 $(N+1)\times m$ 的线性方程组 $\boldsymbol{AL} = \boldsymbol{b}$ 存在的非负解为 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$, 则有 $\boldsymbol{AL}^* = \boldsymbol{b}$, 利用矩阵乘法将其写成分量形式得

$$\sum_{i=1}^{m} e_{it} l_i^* = b_t, \quad t = 1, 2, \cdots, N, \quad \sum_{i=1}^{m} l_i^* = 1.$$

因为 $b_t \in (-|e_{1t}|, |e_{1t}|)$, 所以有

$$\left|\sum_{i=1}^{m} l_i^* e_{it}\right| = |b_t| < |e_{1t}|, \quad t = 1, 2, \cdots, N. \tag{6.3.14}$$

由式 (6.3.10)、(6.3.11)、(6.3.14) 知

$$d(\boldsymbol{L}^*) = \sum_{t=1}^{N} \left|\sum_{i=1}^{m} l_i^* e_{it} \Bigg/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2}\right| < \sum_{t=1}^{N} \left| e_{1t} \Bigg/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2}\right| = d_1.$$

又由条件知 $d_1 = \min\{d_i, i = 1, 2, \cdots, m\}$, 所以 $d(\boldsymbol{L}^*) < \min\{d_i, i = 1, 2, \cdots, m\}$. 由定义 6.3.3 知结论成立. 证毕.

推论 6.3.2 设 $d_1 = \min\{d_i, i = 1, 2, \cdots, m\}$, 若倒数误差信息矩阵 $\boldsymbol{E} = (e_{it})_{m\times N}$ 中任一列 m 个元素 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的算术平均数 $\bar{e}_t = \dfrac{1}{m}\sum_{i=1}^{m} e_{it} \in (-|e_{1t}|, |e_{1t}|), t = 1, 2, \cdots, N$, 则简单等权平均组合预测方法是优性组合预测方法.

6.3.4 冗余单项预测方法的一个判定

定理 6.3.4 在倒数误差信息矩阵 $\boldsymbol{E} = (e_{it})_{m\times N}$中, 任意第 t 列元素 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的符号完全相同, $t = 1, 2, \cdots, N$, 且第 j 种单项预测方法优超第 k 种单项预测方法, 则组合预测 (6.3.12) 式的冗余度至少为 $1/m$. 即至少存在第 k 种单项预测方法为冗余预测方法.

证明 采用反证法. 假设第k种单项预测方法不为冗余预测方法, 即在式 (6.3.12) 的最优解 $\boldsymbol{L}^* = (l_1^*, \cdots, l_j^*, \cdots, l_k^*, \cdots, l_m^*)^{\mathrm{T}}$ 中, 第 k 种单项预测方法对应的最优权系数 $l_k^* > 0$.

因为任意第 t 列元素 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的符号完全相同, $t = 1, 2, \cdots, N$, 并注意到最优解 $\boldsymbol{L}^*$ 满足非负性, 所以 $\left|\sum_{i=1}^{m} l_i^* e_{it}\right| = \sum_{i=1}^{m} l_i^* |e_{it}|, t = 1, 2, \cdots, N$, 从而最优解 $\boldsymbol{L}^*$ 对应的目标函数值为

$$d(\boldsymbol{L}^*) = \sum_{t=1}^{N} \left|\sum_{i=1}^{m} l_i^* e_{it} \Big/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2}\right| = \sum_{t=1}^{N} \sum_{i=1}^{m} l_i^* |e_{it}| \Big/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2}. \tag{6.3.15}$$

令 $\hat{\boldsymbol{L}} = (l_1^*, \cdots, l_j^* + l_k^*, \cdots, 0, \cdots, l_m^*)^{\mathrm{T}}$, 即 $\hat{\boldsymbol{L}}$ 的第 j 个分量为 $l_j^* + l_k^*$, 第 k 个分量为 0, 其他 $m-2$ 个分量与 $\boldsymbol{L}^*$ 相同, 显然 $\hat{\boldsymbol{L}}$ 为式 (6.3.12) 的一个可行解, 同理, $\hat{\boldsymbol{L}}$ 对应的目标函数值为

$$\begin{aligned} d(\hat{\boldsymbol{L}}) &= \sum_{t=1}^{N} \left(\sum_{i \neq j,k} l_i^* |e_{it}| + (l_j^* + l_k^*) |e_{jt}| + 0 \times |e_{kt}|\right) \Big/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2} \\ &= \sum_{t=1}^{N} \left(\sum_{i=1}^{m} l_i^* |e_{it}| + l_k^* (|e_{jt}| - |e_{kt}|)\right) \Big/ \sqrt{\sum_{i=1}^{m} \tilde{x}_{it}^2}. \end{aligned} \tag{6.3.16}$$

因为第 j 种单项预测方法优超第 k 种单项预测方法, 由定义 6.3.4 知

$$|e_{jt}| \leqslant |e_{kt}|, \quad t = 1, 2, \cdots, N, \quad |e_{jt_0}| < |e_{kt_0}|, \quad t_0 \in \{1, 2, \cdots, N\}. \tag{6.3.17}$$

由式 (6.3.15)~(6.3.17) 并注意到假设 $l_k^* > 0$, 所以有

$$d(\hat{\boldsymbol{L}}) < d(\boldsymbol{L}^*).$$

而这与 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$ 为式 (6.3.12) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 即知式 (6.3.12) 的冗余度至少为 $1/m$. 证毕.

6.4 广义加权算术平均组合预测法的最优化理论基础及性质

文献 [39] 提出了一种广义的非线性加权平均组合预测模型, 并给出了 p 次幂平均绝对误差、p 次幂最大误差、p 次幂均方差等准则下的组合预测模型权系数的确定方法. 在文献 [39] 的基础上, 从 p 次幂误差的概念出发, 通过构造最优化函数进一步探讨了该方法的最优化理论依据及其数学性质.

6.4.1 广义加权算术平均组合预测法的最优化理论基础[86]

设某社会经济现象的指标序列的观察值为 $\{x_t, t = 1, 2, \cdots, N\}$, 设有 m 个单项预测方法对其进行预测, x_{it} 为第 i 种预测方法第 t 时刻的预测值, $i = 1, 2, \cdots, m., t =$

$1, 2, \cdots, N$. 由于受到多种因素的影响, 预测误差总是不可避免的, 则有如下几个概念.

定义 6.4.1 称 $e_{it}^{(p)} = x_t^p - x_{it}^p$ 为第 i 种预测方法在第 t 时刻的 p 次幂误差, p 为正常数, $i = 1, 2, \cdots, m$, 称 $\boldsymbol{\varepsilon}_i^{(p)} = (e_{i1}^{(p)}, e_{i2}^{(p)}, \cdots, e_{iN}^{(p)})^{\mathrm{T}}$ 为第 i 种预测方法在各个时刻的 p 次幂误差向量, 称矩阵 $\boldsymbol{E}^{(p)} = (\boldsymbol{\varepsilon}_1^{(p)}, \boldsymbol{\varepsilon}_2^{(p)}, \cdots, \boldsymbol{\varepsilon}_m^{(p)})_{m\times N}^{\mathrm{T}}$ 为组合预测模型的 p 次幂预测误差矩阵.

定义 6.4.2 若组合预测模型采用如下的非线性加权平均形式

$$\hat{x}_t = \left(\sum_{i=1}^{m} \omega_i x_{it}^p\right)^{1/p}. \tag{6.4.1}$$

则称该组合预测方法为广义加权算术平均组合预测方法, 其中 ω_i 表示第 i 种预测方法在组合预测中的权系数, $\sum\limits_{i=1}^{m} \omega_i = 1$, $\omega_i \geqslant 0, i = 1, 2, \cdots, m$. $\hat{x}_t$ 表示第 t 时刻的组合预测值.

特别地, 当 p=1 时, 式 (6.4.1) 就成为 $\hat{x}_t = \sum\limits_{i=1}^{m} \omega_i x_{it}$, 这就是线性的加权平均组合预测方法.

在统计学的回归分析中, 我们经常使用误差平方和作为实际值和预测值的偏离程度. 因此, 对于广义加权算术平均组合预测方法而言, 为了综合利用 m 个单项预测方法所提供的信息, 我们当然希望组合预测方法在第 t 时刻的组合预测值与 m 个单项预测方法预测值的 p 次幂误差平方和加权算术平均达到最小, 这可用来反映组合预测方法和 m 个单项预测方法的逼近程度, 于是构造如下最优化函数

$$\min J(t) = \sum_{i=1}^{m} \omega_i (\hat{x}_t^p - x_{it}^p)^2, \quad t = 1, 2, \cdots, N, \tag{6.4.2}$$

其中 $\hat{x}_t, \omega_i$ 表示的含义同上.

在 m 个单项预测方法的加权系数 $\omega_1, \omega_2, \cdots, \omega_m$ 已知的条件下, 广义加权算术平均组合预测方法究竟是否存在某个准则, 使其在该准则下为最优的. 亦即要探讨一下广义加权算术平均组合预测方法的最优化基础, 下面的定理 6.4.1 给予了回答.

定理 6.4.1 若广义加权算术平均组合预测方法采用式 (6.4.1) 的非线性加权平均形式, 则它正好是最优化函数式 (6.4.2) 的最优解.

证明 由式 (6.4.2) 可知, $J(t)$ 为 $\hat{x}_t$ 的幂函数, 根据极值的必要条件, 令

$$\frac{\mathrm{d}J(t)}{\mathrm{d}\hat{x}_t} = 0, \quad t = 1, 2, \cdots, N.$$

从而有

$$2\sum_{i=1}^{m} \omega_i (\hat{x}_t^p - x_{it}^p)\hat{x}_t^{p-1} = 0.$$

考虑到 $\sum_{i=1}^{m}\omega_i=1$, 所以得唯一驻点

$$\hat{x}_t=\left(\sum_{i=1}^{m}\omega_i x_{it}^p\right)^{1/p}.$$

又因为 $\dfrac{\mathrm{d}^2J(t)}{\mathrm{d}\hat{x}_t^2}=2\hat{x}_t^{p-2}\sum_{i=1}^{m}\omega_i[(2p^2-p)\hat{x}_t^p-(p^2-p)x_{it}^p)]$, 所以在驻点 $\hat{x}_t=\left(\sum_{i=1}^{m}\omega_i x_{it}^p\right)^{1/p}$ 处

$$\frac{\mathrm{d}^2J(t)}{\mathrm{d}\hat{x}_t^2}=2p^2\left(\sum_{i=1}^{m}\omega_i x_{it}^p\right)^{(2p-2)/p}>0.$$

因此驻点 $\hat{x}_t$ 为 $J(t)$ 的最小值点, 即结论成立. 证毕.

6.4.2 广义加权算术平均组合预测法的几个概念

定义 6.4.3 称 $e_{ct}^{(p)}=x_t^p-\hat{x}_t^{(p)}$ 为组合预测方法在第 t 时刻的 p 次幂预测误差, p 为正常数, $i=1,2,\cdots,m$, 称 $e_i^{(p)}=\sum_{t=1}^{N}\left|e_{it}^{(p)}\right|$ 为第 i 种预测方法 p 次幂预测误差绝对值之和, 称 $e_c^{(p)}=\sum_{t=1}^{N}\left|e_{ct}^{(p)}\right|$ 为组合预测方法 p 次幂预测误差绝对值之和.

显然, $e_i^{(p)}$ 和 $e_c^{(p)}$ 分别反映单项预测方法和组合预测方法的预测精度的指标, 它们满足不等式

$$e_c^{(p)}\leqslant\sum_{i=1}^{m}\omega_i e_i^{(p)}. \tag{6.4.3}$$

事实上, 注意到 $\sum_{i=1}^{m}\omega_i=1$, $\omega_i\geqslant 0$, 则有

$$\begin{aligned}e_c^{(p)}&=\sum_{t=1}^{N}\left|e_{ct}^{(p)}\right|=\sum_{t=1}^{N}\left|x_t^p-\hat{x}_t^{(p)}\right|=\sum_{t=1}^{N}\left|x_t^p-\sum_{i=1}^{m}\omega_i x_{it}^p\right|=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}\omega_i e_{it}^{(p)}\right|\\&\leqslant\sum_{t=1}^{N}\sum_{i=1}^{m}\omega_i\left|e_{it}^{(p)}\right|=\sum_{i=1}^{m}\omega_i\sum_{t=1}^{N}\left|e_{it}^{(p)}\right|=\sum_{i=1}^{m}\omega_i e_i^{(p)}.\end{aligned}$$

则以组合预测方法 p 次幂预测误差绝对值之和达到最小的广义加权算术平均组合预测模型可表示成如下模型

$$\min e_c^{(p)}=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}\omega_i e_{it}^{(p)}\right|,$$

$$\text{s.t.}\begin{cases}\sum_{i=1}^{m}\omega_i=1,\\ \omega_i\geqslant 0, i=1,2,\cdots,m.\end{cases} \tag{6.4.4}$$

记 $e_{\min}^{(p)}=\min\limits_{1\leqslant i\leqslant m}\{e_i^{(p)}\}, e_{\max}^{(p)}=\max\limits_{1\leqslant i\leqslant m}\{e_i^{(p)}\}$, 即 $e_{\min}^{(p)}$ 表示 m 种预测方法中的最小 p 次幂预测误差绝对值之和, $e_{\max}^{(p)}$ 表示 m 种预测方法中的最大 p 次幂预测误差绝对值之和.

定义 6.4.4 若 $e_c^{(p)}>e_{\max}^{(p)}$, 则称组合预测模型 (6.4.4) 为劣性组合预测, 若 $e_{\min}^{(p)}\leqslant e_c^{(p)}\leqslant e_{\max}^{(p)}$, 则称组合预测模型 (6.4.4) 为非劣性组合预测, 若 $e_c^{(p)}<e_{\min}^{(p)}$, 则称组合预测模型 (6.4.4) 为优性组合预测.

定义 6.4.4 表明, 只有组合预测 p 次幂预测误差绝对值之和小于各单项预测中最小者, 则该组合预测模型为优性组合预测, 此时组合预测才有存在的意义.

定义 6.4.5 若组合预测模型的 p 次幂预测误差矩阵 $\boldsymbol{E}^{(p)}$ 的第 i 行和第 k 行元素满足不等式 $|e_{it}^{(p)}|\leqslant|e_{kt}^{(p)}|, t=1,2,\cdots,N$, 且至少对某个时刻 t_0 有严格的不等号成立, $t_0\in\{1,2,\cdots,N\}$. 则称第 i 种单项预测方法优超第 k 种单项预测方法.

定义 6.4.5 表明, 第 i 种单项预测方法在各个时刻 p 次幂预测误差的绝对值不大于第 k 种单项预测方法, 而且至少对某个时刻 t_0 处 p 次幂预测误差的绝对值严格小于第 k 种单项预测方法, 直观上可以认为第 i 种单项预测方法要 “好于” 第 k 种单项预测方法.

6.4.3 广义加权算术平均组合预测法的数学性质

定理 6.4.2 组合预测模型 (6.4.4) 的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 设 $\boldsymbol{\Omega}=(\omega_1,\omega_2,\cdots,\omega_m)^{\mathrm{T}}$ 为组合预测模型 (6.4.4) 的任一个可行解, 则有

$$\sum_{i=1}^{m}\omega_i=1,\quad \omega_i\geqslant 0,\quad i=1,2,\cdots,m.$$

由式 (6.4.3) 知, $e_c^{(p)}\leqslant\sum\limits_{i=1}^{m}\omega_i e_i^{(p)}\leqslant e_{\max}^{(p)}$, 再由定义 6.4.4 得结论成立. 证毕.

定理 6.4.2 表明, 从组合预测 p 次幂预测误差的绝对值之和这个角度而言, 组合预测模型 (6.4.4) 的任一个归一化非负权系数所对应的组合预测均不会比 “最差” 的单项预测方法还要 “差”.

推论 6.4.1 简单平均组合预测方法至少是非劣性组合预测.

定理 6.4.3 设 $e_1^{(p)}=\min\limits_{1\leqslant i\leqslant m}\{e_i^{(p)}\}, b_t\in(-|e_1^{(p)}|,|e_1^{(p)}|), t=1,2,\cdots,N$, 若如下 $(N+1)\times m$ 的线性方程组

$$\begin{cases} e_{11}^{(p)}\omega_1 + e_{21}^{(p)}\omega_2 + \cdots + e_{m1}^{(p)}\omega_m = b_1, \\ e_{12}^{(p)}\omega_1 + e_{22}^{(p)}\omega_2 + \cdots + e_{m2}^{(p)}\omega_m = b_2, \\ \qquad\cdots\cdots\cdots\cdots \\ e_{1N}^{(p)}\omega_1 + e_{2N}^{(p)}\omega_2 + \cdots + e_{mN}^{(p)}\omega_m = b_N, \\ \omega_1 + \omega_2 + \cdots + \omega_m = 1 \end{cases} \tag{6.4.5}$$

存在非负解, 则组合预测模型 (6.4.4) 一定存在优性组合预测.

证明 设如上的 $(N+1)\times m$ 的线性方程组存在的非负解为 $\boldsymbol{\Omega}^* = (\omega_1^*, \omega_2^*, \cdots, \omega_m^*)^{\mathrm{T}}$, 则有

$$\sum_{i=1}^{m} e_{it}^{(p)}\omega_i^* = b_t, \quad t = 1, 2, \cdots, N, \quad \sum_{i=1}^{m}\omega_i^* = 1, \quad \omega_i^* \geqslant 0.$$

因为 $b_t \in (-|e_{1t}^{(p)}|, |e_{1t}^{(p)}|), t = 1, 2, \cdots, N$, 所以有

$$\left|\sum_{i=1}^{m} e_{it}^{(p)}\omega_i^*\right| < \left|e_{1t}^{(p)}\right|, \quad t = 1, 2, \cdots, N.$$

因而 $\sum_{t=1}^{N}\sum_{i=1}^{m}|\omega_i^* e_{it}^{(p)}| < \sum_{t=1}^{N}|e_{1t}^{(p)}|$, 即 $e_c^{(p)} < e_1^{(p)} = e_{\min}^{(p)}$, 由定义 6.4.4 知结论成立. 证毕.

推论 6.4.2 设 $e_1^{(p)} = \min\limits_{1\leqslant i\leqslant m}\{e_i^{(p)}\}$, 即 $e_1^{(p)}$ 表示 m 种单项预测方法中的最小 p 次幂预测误差, 若组合预测模型的最小 p 次幂预测误差矩阵 $\boldsymbol{E}^{(p)} = \left(e_{it}^{(p)}\right)_{m\times N}$ 中任一列 m 个元素 $e_{1t}^{(p)}, e_{2t}^{(p)}, \cdots, e_{mt}^{(p)}$ 的算术平均数 $\bar{e}_t^{(p)} = \dfrac{1}{m}\sum\limits_{i=1}^{m} e_{it}^{(p)}$ 且 $\bar{e}_t^{(p)} \in \left(-|e_{1t}^{(p)}|, |e_{1t}^{(p)}|\right)$, $t = 1, 2, \cdots, N$, 则简单平均组合预测方法是优性组合预测方法.

证明 设 $\boldsymbol{\Omega} = (1/m, 1/m, \cdots, 1/m)^{\mathrm{T}}$ 为简单平均组合预测方法的权系数, 由条件知

$$\bar{e}_t^{(p)} = \frac{1}{m}\sum_{i=1}^{m} e_{it}^{(p)} \in \left(-|e_{1t}^{(p)}|, |e_{1t}^{(p)}|\right), \quad t = 1, 2, \cdots, N.$$

从而 $\boldsymbol{\Omega} = (1/m, 1/m, \cdots, 1/m)^{\mathrm{T}}$ 是满足于线性方程组 (6.4.5) 的非负解, 从而由定理 6.4.3 得推论成立. 证毕.

对于组合预测模型 (6.4.4) 有如下结论成立.

定理 6.4.4 组合预测模型 (6.4.4) 的最优目标函数是 m 的单调不增函数, 即

$$e_c^{(p)*}(\omega_1, \omega_2, \cdots, \omega_m) \geqslant e_c^{(p)*}(\omega_1, \omega_2, \cdots, \omega_{m+1}),$$

其中 m 为参与组合预测的各单项预测方法总个数, $e_c^{(p)*}(\omega_1,\omega_2,\cdots,\omega_m)$ 表示 m 个单项预测方法参与的组合预测模型(6.4.4)对应的最优目标函数, $e_c^{(p)*}(\omega_1,\omega_2,\cdots,\omega_{m+1})$ 表示再增加一个单项预测方法参与的组合预测模型 (6.4.4) 对应的最优目标函数.

证明 设 $\boldsymbol{\Omega}^*=(\omega_1^*,\omega_2^*,\cdots,\omega_m^*)^{\mathrm{T}}$ 为 m 个单项预测方法参与的组合预测模型 (6.4.4) 的最优解, 所以

$$e_c^{(p)*}(\omega_1,\omega_2,\cdots,\omega_m)=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}\omega_i^* e_{it}^{(p)}\right|,$$

其中 $\sum_{i=1}^{m}\omega_i^*=1,\omega_i^*\geqslant 0,i=1,2,\cdots,m$. 同理, 设 $\bar{\boldsymbol{\Omega}}=(\bar{\omega}_1,\bar{\omega}_2,\cdots,\bar{\omega}_m,\bar{\omega}_{m+1})^{\mathrm{T}}$ 为再增加一个单项预测方法, 共 $m+1$ 个单项预测方法参与的组合预测模型 (6.4.4) 的最优解, 则有

$$e_c^{(p)*}(\omega_1,\omega_2,\cdots,\omega_{m+1})=\sum_{t=1}^{N}\left|\sum_{i=1}^{m+1}\bar{\omega}_i e_{it}^{(p)}\right|,$$

其中 $\sum_{i=1}^{m+1}\bar{\omega}_i=1,\bar{\omega}_i\geqslant 0,i=1,2,\cdots,m+1$. 令 $\boldsymbol{\Omega}=(\omega_1^*,\omega_2^*,\cdots,\omega_m^*,0)^{\mathrm{T}}$, 显然, $\boldsymbol{\Omega}$ 为 $m+1$ 个单项预测方法参与的组合预测模型 (6.4.4) 的可行解, 则有

$$\begin{aligned}e_c^{(p)*}(\omega_1,\omega_2,\cdots,\omega_{m+1})&\leqslant e_c^{(p)}(\omega_1^*,\omega_2^*,\cdots,\omega_m^*,0)\\&=\sum_{t=1}^{N}\left|\sum_{i=1}^{m}\omega_i^* e_{it}^{(p)}\right|=e_c^{(p)*}(\omega_1,\omega_2,\cdots,\omega_m).\end{aligned}$$

从而结论成立. 证毕.

定理 6.4.4 证明, 当再增加一个单项预测方法时, 组合预测模型 (6.4.4) 的最优目标函数值可能不变. 这表明组合预测模型 (6.4.4) 中可能存在冗余预测方法. 下面两个定理为冗余信息提供了判定.

定理 6.4.5 若组合预测模型的 p 次幂预测误差矩阵 $\boldsymbol{E}^{(p)}=\left(e_{it}^{(p)}\right)_{m\times N}$ 中, 任意第 t 列元素 $e_{1t}^{(p)},e_{2t}^{(p)},\cdots,e_{mt}^{(p)}$ 的符号完全相同, $t=1,2,3,\cdots,N$. 设 $e_1^{(p)}=\min\limits_{1\leqslant i\leqslant m}\{e_i^{(p)}\}$, 则组合预测模型 (6.4.4) 的冗余度为 $(m-1)/m$, 即后面 $m-1$ 种单项预测方法均为冗余预测方法.

证明 因为 $\boldsymbol{E}^{(p)}$ 中任意第 t 列元素 $e_{1t}^{(p)},e_{2t}^{(p)},\cdots,e_{mt}^{(p)}$ 的符号完全相同, $t\in\{1,2,\cdots,N\}$, 所以有 $\left|\sum_{i=1}^{m}\omega_i e_{it}^{(p)}\right|=\sum_{i=1}^{m}\omega_i\left|e_{it}^{(p)}\right|,t\in\{1,2,\cdots,N\}$.

注意到 $\sum_{i=1}^{m}\omega_i=1,\omega_i\geqslant 0$, 则有

$$e_c^{(p)} = \sum_{t=1}^{N}\left|\sum_{i=1}^{m}\omega_i e_{it}^{(p)}\right| = \sum_{t=1}^{N}\sum_{i=1}^{m}\omega_i\left|e_{it}^{(p)}\right| = \sum_{i=1}^{m}\omega_i\sum_{t=1}^{N}\left|e_{it}^{(p)}\right| = \sum_{i=1}^{m}\omega_i e_i^{(p)}.$$

因为 $e_1^{(p)} = \min\limits_{1\leqslant i\leqslant m}\{e_i^{(p)}\}$, 所以 $e_c^{(p)} = \sum\limits_{i=1}^{m}\omega_i e_i^{(p)} \geqslant e_1^{(p)}$, 等号成立当且仅当 $\boldsymbol{\Omega}^* = (\omega_1^*, \omega_2^*, \cdots, \omega_m^*)^{\mathrm{T}} = (1, 0, \cdots, 0)^{\mathrm{T}}$, 即组合预测模型 (6.4.4) 的最优解为 $(1, 0, \cdots, 0)^{\mathrm{T}}$, 从而后面 $m-1$ 种单项预测方法均为冗余预测方法. 证毕.

定理 6.4.6　若组合预测模型的 p 次幂预测误差矩阵 $\boldsymbol{E}^{(p)} = \left(e_{it}^{(p)}\right)_{m\times N}$ 中, 任意第 t 列元素 $e_{1t}^{(p)}, e_{2t}^{(p)}, \cdots, e_{mt}^{(p)}$ 的符号完全相同, $t = 1, 2, \cdots, N$, 且第 i 种单项预测方法优超第 k 种单项预测方法, 则组合预测模型 (6.4.4) 的冗余度至少为 $1/m$. 即至少存在一种单项预测方法为冗余预测方法.

证明　因为 $\boldsymbol{E}^{(p)}$ 中任意第 t 列元素 $e_{1t}^{(p)}, e_{2t}^{(p)}, \cdots, e_{mt}^{(p)}$ 的符号完全相同, $t \in \{1, 2, \cdots, N\}$, 在定理 6.4.5 中已证明得

$$e_c^{(p)} = \sum_{i=1}^{m}\omega_i e_i^{(p)}.$$

又因为第 i 种单项预测方法优超第 k 种单项预测方法, 所以由定义 6.4.5 知 $\left|e_{it}^{(p)}\right| \leqslant \left|e_{kt}^{(p)}\right|, t = 1, 2, \cdots, N$, 且至少对某个 t_0 成立 $\left|e_{it_0}^{(p)}\right| < \left|e_{kt_0}^{(p)}\right|, t_0 \in \{1, 2, \cdots, N\}$, 所以有

$$e_i^{(p)} = \sum_{t=1}^{N}\left|e_{it}^{(p)}\right| < \sum_{t=1}^{N}\left|e_{kt}^{(p)}\right| = e_k^{(p)}, \quad \text{即} e_i^{(p)} < e_k^{(p)}.$$

假设第 k 种单项预测方法不为冗余预测方法, 设组合预测模型 (6.4.4) 的最优解为

$$\boldsymbol{\Omega} = (\omega_1^*, \cdots, \omega_i^*, \cdots, \omega_k^*, \cdots, \omega_m^*)^{\mathrm{T}}, \quad \omega_k^* \neq 0. \quad \sum_{i=1}^{m}\omega_i^* = 1.$$

$\boldsymbol{\Omega}$ 对应的目标函数值为

$$e_c^{(p)} = \omega_1^* e_1^{(p)} + \cdots + \omega_i^* e_i^{(p)} + \cdots + \omega_k^* e_k^{(p)} + \cdots + \omega_m^* e_m^{(p)}.$$

构造组合预测模型 (6.4.1) 的另一可行解为 $\hat{\boldsymbol{\Omega}} = (\omega_1^*, \cdots, \omega_i^* + \omega_k^*, \cdots, 0, \cdots, \omega_m^*)^{\mathrm{T}}$, $\hat{\boldsymbol{\Omega}}$ 对应的目标函数值为

$$\hat{e}_c^{(p)} = \omega_1^* e_1^{(p)} + \cdots + (\omega_i^* + \omega_k^*) e_i^{(p)} + \cdots + 0 \times e_k^{(p)} + \cdots + \omega_m^* e_m^{(p)}.$$

由 $e_i^{(p)} < e_k^{(p)}$ 知

$$\hat{e}_c^{(p)} < e_c^{(p)}.$$

而这与 $\boldsymbol{\Omega} = (\omega_1^*, \cdots, \omega_i^*, \cdots, \omega_k^*, \cdots, \omega_m^*)^{\mathrm{T}}$ 为组合预测模型 (6.4.4) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 此即组合预测模型 (6.4.4) 的冗余度至少为 $1/m$. 证毕.

第 7 章　基于相关性指标的最优组合预测的有效性理论

7.1　基于相关系数的优性组合预测模型的性质

目前传统的组合预测方法均是直接从改善某种拟合误差角度提出来的[2~4,18,42,87,88]. 组合预测模型根据建立的某个准则的优劣程度，可以分为劣性组合预测、非劣性组合预测、优性组合预测. 文献 [18] 针对无非负约束的以误差平方和达到最小的组合预测模型，提出了劣性组合预测、非劣性组合预测、优性组合预测的概念，并利用组合预测绝对误差信息矩阵的性质判断简单算术平均组合预测方法是非劣性组合预测、优性组合预测的条件. 文献 [72] 提出了基于预测有效度的组合预测模型，并给出组合预测权系数的线性规划的求解方法. 文献 [73] 针对基于预测有效度的组合预测模型提出了优性组合预测、预测方法优超和冗余度的概念，探讨了优性组合预测存在的充分条件，也给出冗余信息出现的判定. 文献 [74] 研究了考虑预测精度标准差的预测有效度的组合预测模型的性质.

文献 [46] 给出了研究组合预测方法的另一新途径，即提出了基于相关性指标的最优组合预测模型的研究，与传统的组合预测方法有较大的差别. 然而，文献 [46] 只是从实证分析角度说明了基于相关性的组合预测方法也能够取得比较好的组合预测效果. 本节在此基础上针对基于相关系数的组合预测模型，给出新的优性组合预测、预测方法优超等概念，研究了非劣性组合预测、优性组合预测和冗余预测方法的存在性等基本问题，并给出冗余预测方法的判定，从而在理论上说明基于相关系数的组合预测方法的有效性.

7.1.1　组合预测协方差信息矩阵性质及模型[90]

设 $\hat{x}_t=\sum\limits_{i=1}^{m}l_ix_{it}$ 为实际观察值 x_t 的组合预测值，$l_1,l_2,\cdots,l_m$ 为 m 种单项预测方法的加权系数，且满足 $\sum\limits_{i=1}^{m}l_i=1$, $l_i\geqslant 0$, $i=1,2,\cdots,m$, 其中指标序列的实际观察值为$\{x_t, t=1,2,\cdots,N\}$, $\{x_{it}, t=1,2,\cdots,N\}$为第 i 种预测方法进行预测的预测值序列, i=1,2,$\cdots$, m. 则有如下几个概念.

定义 7.1.1　称 $e_{it}=x_{it}-\bar{x}_i$ 为第 i 种预测方法预测值对其算术平均数在第 t

时刻的离差, $i=1,2,\cdots,m$, $t=1,2,\cdots,N$. $e_t=x_t-\bar{x}$ 指标序列的实际观察值对其算术平均数在第 t 时刻的离差, $\hat{e}_t=\hat{x}_t-\bar{\hat{x}}$ 为组合预测值与其算术平均数在第 t 时刻的离差. 其中 $\bar{x}_i=\frac{1}{N}\sum_{t=1}^{N}x_{it}$, $\bar{x}=\frac{1}{N}\sum_{t=1}^{N}x_t$, $\bar{\hat{x}}=\frac{1}{N}\sum_{t=1}^{N}\hat{x}_t$ 为相应的算术平均数.

对于 t 时刻组合预测值的离差和第 i 种预测方法预测值的离差满足如下关系

$$\hat{e}_t=\sum_{i=1}^{m}l_ie_{it}.$$

实际上有

$$\begin{aligned}\hat{e}_t=\hat{x}_t-\bar{\hat{x}}&=\sum_{i=1}^{m}l_ix_{it}-\frac{1}{N}\sum_{t=1}^{N}\hat{x}_t\\&=\sum_{i=1}^{m}l_ix_{it}-\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}l_ix_{it}=\sum_{i=1}^{m}l_i(x_{it}-\bar{x}_i)=\sum_{i=1}^{m}l_ie_{it}.\end{aligned}$$

定义 7.1.2　令

$$R_i=\frac{\sum_{t=1}^{N}(x_t-\bar{x})(x_{it}-\bar{x}_i)}{\sqrt{\sum_{t=1}^{N}(x_t-\bar{x})^2}\sqrt{\sum_{t=1}^{N}(x_{it}-\bar{x}_i)^2}},\quad i=1,2,\cdots,m,$$

$$R=\frac{\sum_{t=1}^{N}(x_t-\bar{x})(\hat{x}_t-\bar{\hat{x}})}{\sqrt{\sum_{t=1}^{N}(x_t-\bar{x})^2}\sqrt{\sum_{t=1}^{N}(\hat{x}_t-\bar{\hat{x}})^2}}.$$

则称 R_i 为第 i 种单项预测方法预测值序列与实际观察值序列的相关系数, 称 R 为组合预测值序列与实际观察值序列的相关系数.

相关系数 R_i, R 也可以用离差序列来表达, 则有

$$R_i=\frac{\sum_{t=1}^{N}e_te_{it}}{\sqrt{\sum_{t=1}^{N}e_t^2}\sqrt{\sum_{t=1}^{N}e_{it}^2}},\quad R=\frac{\sum_{t=1}^{N}e_t\sum_{i=1}^{m}l_ie_{it}}{\sqrt{\sum_{t=1}^{N}e_t^2}\sqrt{\sum_{t=1}^{N}\left(\sum_{i=1}^{m}l_ie_{it}\right)^2}}.\tag{7.1.1}$$

记 $\boldsymbol{e}_i = (e_{i1}, e_{i2}, \cdots, e_{iN})^{\mathrm{T}}$ 为第 i 种单项预测方法的预测离差列向量, 令 $E_{ij} = \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}_j = \sum\limits_{t=1}^{N} e_{it} e_{jt}$, $i, j = 1, 2, \cdots, m$, $\boldsymbol{E} = (E_{ij})_{m \times m}$, 则$\boldsymbol{E}$称为组合预测协方差信息矩阵. 当 $i \neq j$ 时, E_{ij} 表示第 i 种与第 j 种单项预测方法预测值序列的协方差, 当 $i = j$ 时, E_{ii} 表示第 i 种单项预测方法预测值序列的方差. 对于组合预测协方差信息矩阵$\boldsymbol{E}$有如下结论.

定理 7.1.1 若 m $(m < N)$ 种单项预测方法的预测离差向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_m$ 是线性无关的, 则 $\boldsymbol{E}$ 为正定矩阵. 其中 $\boldsymbol{e}_i = (e_{i1}, e_{i2}, \cdots, e_{iN})^{\mathrm{T}}$, $i = 1, 2, \cdots, m$.

证明与定理 4.2.1 类似, 故省略.

记 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为组合预测加权系数列向量, 因为 $\sum\limits_{t=1}^{N} \left(\sum\limits_{i=1}^{m} l_i e_{it} \right)^2 = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{m} l_i l_j \left(\sum\limits_{t=1}^{N} e_{it} e_{jt} \right) = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{m} l_i l_j E_{ij} = \boldsymbol{L}^{\mathrm{T}} \boldsymbol{E} \boldsymbol{L}$. 所以 (7.1.1) 式可写成

$$R_i = \frac{\sum\limits_{t=1}^{N} e_t e_{it}}{\sqrt{\sum\limits_{t=1}^{N} e_t^2} \sqrt{E_{ii}}}, \qquad R = \frac{\sum\limits_{i=1}^{m} l_i \sum\limits_{t=1}^{N} e_t e_{it}}{\sqrt{\sum\limits_{t=1}^{N} e_t^2} \sqrt{\boldsymbol{L}^{\mathrm{T}} \boldsymbol{E} \boldsymbol{L}}}. \tag{7.1.2}$$

显然, 组合预测值序列与实际观察值序列的相关系数 R 为各种单项预测方法的加权系数 $l_1, l_2, \cdots, l_m$ 的函数, 记为 $R(l_1, l_2, \cdots, l_m)$. 当从相关系数角度考察组合预测问题的时候, 我们希望 $R(l_1, l_2, \cdots, l_m)$ 愈大愈好, $R(l_1, l_2, \cdots, l_m)$ 越大表示组合预测方法越有效. 当组合预测值序列与实际观察值序列完全相同时, 相关系数达到了最大值 1. 然而预测误差是不可避免的. 因此, 基于相关系数的组合预测模型可表示成如下最优化模型[43]

$$\max R(l_1, l_2, \cdots, l_m) = \frac{\sum\limits_{i=1}^{m} l_i \sum\limits_{t=1}^{N} e_t e_{it}}{\sqrt{\sum\limits_{t=1}^{N} e_t^2} \sqrt{\boldsymbol{L}^{\mathrm{T}} \boldsymbol{E} \boldsymbol{L}}}$$

$$\text{s.t.} \begin{cases} \sum\limits_{i=1}^{m} l_i = 1, \\ l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{7.1.3}$$

7.1.2 基于相关系数的非劣性组合预测和优性组合预测的存在性[90]

记 $R_{\min} = \min\{R_i, i = 1, 2, \cdots, m\}$, $R_{\max} = \max\{R_i, i = 1, 2, \cdots, m\}$.

定义 7.1.3 设 $R(l_1,l_2,\cdots,l_m)$ 为组合预测值序列与实际观察值序列的相关系数, 若 $R(l_1,l_2,\cdots,l_m)<R_{\min}$, 则称权系数 $l_1,l_2,\cdots,l_m$ 确定的组合预测模型为劣性组合预测, 若 $R_{\min}\leqslant R(l_1,l_2,\cdots,l_m)\leqslant R_{\max}$, 则称之为非劣性组合预测, 若 $R(l_1,l_2,\cdots,l_m)>R_{\max}$, 则称之为优性组合预测.

定义 7.1.3 表明, 只有组合预测值序列与预测对象实际观察值序列的相关系数大于各单项预测值序列与实际观察值序列的相关系数中最大者, 则该组合预测模型才是优性的.

定义 7.1.4 若第 i 种、第 k 种预测方法与其他单项预测方法预测值序列的协方差以及它们与实际观察值序列的协方差满足关系式

$$E_{ij}<E_{kj},\quad j=1,2,\cdots,m,\quad \sum_{t=1}^{N}e_te_{it}>\sum_{t=1}^{N}e_te_{kt}. \tag{7.1.4}$$

则称第 i 种单项预测方法基于相关系数优超第 k 种单项预测方法.

实际上, 由定义 7.1.4 知, 当 $j=i$ 时, $E_{ii}<E_{ki}$, 当 $j=k$ 时, $E_{ik}<E_{kk}$, 因为$\boldsymbol{E}$是 m 阶对称阵, 所以 $E_{ik}=E_{ki}$, 从而有 $E_{ii}<E_{kk}$, 因此若第 i 种单项预测方法优超第 k 种, 则由式 (7.1.2)、(7.1.4) 知 $R_i>R_k$. 因此就相关系数而言, 直观上可以认为第 i 种单项预测方法要“好于”第 k 种单项预测方法.

定理 7.1.2 假定 $m\,(m<N)$ 种单项预测方法的预测离差向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 基于相关系数的组合预测模型 (7.1.3) 的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 设 $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为组合预测模型 (7.1.3) 的任一个可行解, 则有

$$\sum_{i=1}^{m}l_i=1,\quad l_i\geqslant 0,\quad i=1,2,\cdots,m. \tag{7.1.5}$$

不妨设 m 种单项预测方法预测值序列与预测对象实际观察值序列的相关系数按大小排序为

$$R_1\geqslant R_2\geqslant\cdots\geqslant R_m.$$

即 $R_1=\max\{R_i,i=1,2,\cdots,m\}$, $R_m=\min\{R_i,i=1,2,\cdots,m\}$, 所以对任意的 R_i, 均有

$$R_i\geqslant R_m,\quad i=1,2,\cdots,m-1.$$

根据定义 7.1.2 得

$$\sum_{t=1}^{N}e_te_{it}\geqslant R_m\sqrt{\sum_{t=1}^{N}e_t^2}\sqrt{E_{ii}},\quad i=1,2,\cdots,m-1.$$

从而有

$$R=\frac{\sum_{i=1}^{m}l_i\sum_{t=1}^{N}e_te_{it}}{\sqrt{\sum_{t=1}^{N}e_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}}}\geqslant\frac{R_m\sum_{i=1}^{m}l_i\sqrt{\sum_{t=1}^{N}e_t^2}\sqrt{E_{ii}}}{\sqrt{\sum_{t=1}^{N}e_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}}}=\frac{R_m\sum_{i=1}^{m}l_i\sqrt{E_{ii}}}{\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}}}. \tag{7.1.6}$$

因为向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 由定理 7.1.1 知组合预测误差信息矩阵$\boldsymbol{E}$为正定阵, 所以它的任意二阶子式 $\begin{vmatrix} E_{ii} & E_{ij} \\ E_{ij} & E_{jj} \end{vmatrix}>0$, 即 $E_{ij}^2<E_{ii}E_{jj}$, 注意到 $l_i\geqslant 0$, $i=1,2,\cdots,m$, 则有

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}=\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_jE_{ij}\leqslant\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_j\sqrt{E_{ii}}\sqrt{E_{jj}}=\left(\sum_{i=1}^{m}l_i\sqrt{E_{ii}}\right)^2. \tag{7.1.7}$$

由式 (7.1.5)、(7.1.6)、(7.1.7) 得

$$R\geqslant R_m=\min\{R_i,i=1,2,\cdots,m\}.$$

再由定义 7.1.4 得结论成立.

定理 7.1.2 表明, 从组合预测相关系数的角度而言, 组合预测模型 (7.1.3) 的任一个归一化非负权系数所对应的组合预测均不会比 "最差" 的单项预测方法还要差.

推论 7.1.1 假定 $m\,(m<N)$ 种单项预测方法的预测离差向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 则简单平均组合预测方法至少是非劣性组合预测.

定义 7.1.5 若某种单项预测方法在组合预测模型最优权系数中为零, 则称该单项预测方法为冗余预测方法. 即该种单项预测方法增加到组合预测模型中不能增加组合预测相关系数, 表明该种单项预测方法只提供冗余信息.

定理 7.1.3 假定 $m\,(m<N)$ 种单项预测方法的预测离差向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 基于相关系数的组合预测模型的冗余度 $k\neq(m-1)/m$, 则其最优解对应的组合预测一定是优性组合预测.

证明 设$\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$ 为模型 (7.1.3) 的最优解, $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为其某一个可行解而不是最优解, 则有

$$\sum_{i=1}^{m}l_i^*=1,\quad l_i^*\geqslant 0,\quad \sum_{i=1}^{m}l_i=1,\quad l_i\geqslant 0,\quad i=1,2,\cdots,m. \tag{7.1.8}$$

组合预测模型 (7.1.3) 目标函数是求最大值的, 则有

$$R(l_1^*,l_2^*,\cdots,l_m^*)>R(l_1,l_2,\cdots,l_m). \tag{7.1.9}$$

由式 (7.1.7)、(7.1.8) 知

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L} \leqslant \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \sqrt{E_{ii}}\sqrt{E_{jj}} \leqslant \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \frac{E_{ii}+E_{jj}}{2} = \sum_{i=1}^{m} l_i E_{ii}.$$

所以有

$$R(l_1,l_2,\cdots,l_m) = \frac{\sum_{i=1}^{m} l_i \sum_{t=1}^{N} e_t e_{it}}{\sqrt{\sum_{t=1}^{N} e_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}}} \geqslant \frac{\sum_{i=1}^{m} l_i \sum_{t=1}^{N} e_t e_{it}}{\sqrt{\sum_{t=1}^{N} e_t^2}\sqrt{\sum_{i=1}^{m} l_i E_{ii}}}. \tag{7.1.10}$$

因为组合预测模型的冗余度 $k \neq (m-1)/m$, 所以最优解中至少有两个非 0 分量, 从而 $(1,0,\cdots,0)^{\mathrm{T}}$, $(0,1,\cdots,0)^{\mathrm{T}}$, $\cdots$, $(0,0,\cdots,1)^{\mathrm{T}}$ 这 m 个 m 维单位列向量分别是组合预测模型 (7.1.3) 的可行解而不是最优解, 由式 (7.1.9)、(7.1.10) 知

$$\begin{cases} R(l_1^*,l_2^*,\cdots,l_m^*) > R(1,0,\cdots,0) \geqslant R_1, \\ R(l_1^*,l_2^*,\cdots,l_m^*) > R(0,1,\cdots,0) \geqslant R_2, \\ \cdots\cdots\cdots\cdots \\ R(l_1^*,l_2^*,\cdots,l_m^*) > R(0,0,\cdots,1) \geqslant R_m. \end{cases}$$

所以

$$R(l_1^*,l_2^*,\cdots,l_m^*) > \max\{R_i, i=1,2,\cdots,m\} = R_{\max}. \tag{7.1.11}$$

再由定义 7.1.4 得结论成立.

定理 7.1.3 表明, 在 m 种单项预测方法参与的基于相关系数的组合预测模型中, 最优解只要有两个以上的分量不为 0, 即两个以上单项预测方法提供有效信息, 则最优解所对应的组合预测方法要 "好于" 单项预测方法中的 "最好" 者, 可见, 最优组合预测模型确实综合利用单项预测方法提供的信息.

在文献 [46] 的实证分析中, 例 1 有两种单项预测方法参与组合预测, 基于相关系数的算术平均组合预测模型的最优解为 $l_1^* = 0.5400$, $l_2^* = 0.4600$. 这表明组合预测模型的冗余度 $k=0$, 且计算出

$$R_1{=}0.9832, \quad R_2{=}0.9801, \quad R(l_1^*,l_2^*) = 0.9902.$$

显然 $R(l_1^*,l_2^*) > \max\{R_1,R_2\}$, 该最优解对应的组合预测一定是优性组合预测. 文献 [46] 的例 2 也有同样的结论. 文献 [46] 的实证分析验证定理 7.1.3 结论.

7.1.3 冗余预测方法的存在性及其判定

定理 7.1.4 基于相关系数的组合预测模型的最优目标函数值是参与组合预测的各单项预测方法总个数 m 的单调不减函数. 即

$$R(l_1^*,l_2^*,\cdots,l_m^*) \leqslant R(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1}),$$

其中 $R(l_1^*, l_2^*, \cdots, l_m^*)$ 表示 m 个单项预测方法参与的组合预测模型 (7.1.3) 对应的最大相关系数,$R(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1})$ 表示再增加一个单项预测方法共 $m+1$ 个单项预测方法对应的最大相关系数.

证明 设 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$ 为 m 个单项预测方法参与的组合预测模型 (7.1.3) 的最优解, 其最优目标函数值为

$$R(l_1^*, l_2^*, \cdots, l_m^*) = \frac{\sum_{i=1}^{m} l_i^* \sum_{t=1}^{N} e_t e_{it}}{\sqrt{\sum_{t=1}^{N} e_t^2 \sqrt{\boldsymbol{L}^{*\mathrm{T}} \boldsymbol{E} \boldsymbol{L}^*}}}, \tag{7.1.12}$$

其中 $\sum_{i=1}^{m} l_i^* = 1$, $l_i^* \geqslant 0$, $i = 1, 2, \cdots, m$.

同理, 设 $\bar{\boldsymbol{L}} = (\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1})^{\mathrm{T}}$ 为再增加一个单项预测方法, 共 $m+1$ 个单项预测方法参与的组合预测模型 (7.1.3) 的最优解, 其最优目标函数值为

$$R(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1}) = \frac{\sum_{i=1}^{m+1} \bar{l}_i \sum_{t=1}^{N} e_t e_{it}}{\sqrt{\sum_{t=1}^{N} e_t^2 \sqrt{\bar{\boldsymbol{L}}^{\mathrm{T}} \boldsymbol{E}_{m+1} \bar{\boldsymbol{L}}}}}, \tag{7.1.13}$$

其中 $\sum_{i=1}^{m+1} \bar{l}_i = 1$, $\bar{l}_i \geqslant 0$, $i = 1, 2, \cdots, m+1$, $\boldsymbol{E}_{m+1}$ 为 $m+1$ 个单项预测方法参与的组合预测协方差信息矩阵, $\boldsymbol{E}_{m+1} = \begin{pmatrix} \boldsymbol{E} & \boldsymbol{a} \\ \boldsymbol{a}^{\mathrm{T}} & E_{(m+1)(m+1)} \end{pmatrix}$, $\boldsymbol{E}$为 m 个单项预测方法参与的组合预测协方差信息矩阵, $\boldsymbol{a} = \left(E_{1(m+1)}, E_{2(m+1)}, \cdots, E_{m(m+1)}\right)^{\mathrm{T}}$ 为前面 m 个单项预测方法与第 $(m+1)$ 个单项预测方法预测序列的协方差向量, 即 $E_{i(m+1)} = \sum_{t=1}^{N} e_{it} e_{(m+1)t}$, $E_{(m+1)(m+1)}$ 表示第 $(m+1)$ 个单项预测方法预测值序列的方差. e_{it} 是由定义 7.1.1 表示的第 i 种预测方法预测值在第 t 时刻的离差, $i = 1, 2, \cdots, m, m+1$.

令 $\boldsymbol{L} = (\boldsymbol{L}^{*\mathrm{T}}, 0)^{\mathrm{T}} = (l_1^*, l_2^*, \cdots, l_m^*, 0)^{\mathrm{T}}$, 显然, $\boldsymbol{L}$ 为 $m+1$ 个单项预测方法参与的组合预测模型 (7.1.3) 的可行解, 而组合预测模型 (7.1.3) 目标函数是求最大值的, 则有

$$R(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1}) \geqslant R(l_1^*, l_2^*, \cdots, l_m^*, 0), \tag{7.1.14}$$

因为 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}_{m+1}\boldsymbol{L}=(\boldsymbol{L}^{*\mathrm{T}},0)\begin{pmatrix}\boldsymbol{E} & \boldsymbol{a}\\ \boldsymbol{a} & E_{(m+1)(m+1)}\end{pmatrix}\begin{pmatrix}\boldsymbol{L}^{*}\\ 0\end{pmatrix}=\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^{*}$, 并注意到式 (7.1.12) , 所以有

$$R(l_1^*,l_2^*,\cdots,l_m^*,0)=\frac{\sum\limits_{i=1}^{m}l_i^*\sum\limits_{t=1}^{N}e_te_{it}+0\times\sum\limits_{t=1}^{N}e_te_{(m+1)t}}{\sqrt{\sum\limits_{t=1}^{N}e_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}_{m+1}\boldsymbol{L}}}$$

$$=\frac{\sum\limits_{i=1}^{m}l_i^*\sum\limits_{t=1}^{N}e_te_{it}}{\sqrt{\sum\limits_{t=1}^{N}e_t^2}\sqrt{\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^{*}}}=R(l_1^*,l_2^*,\cdots,l_m^*). \tag{7.1.15}$$

由式 (7.1.14)、(7.1.15) 得

$$R(l_1^*,l_2^*,\cdots,l_m^*)\leqslant R(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1}).$$

证毕.

定理 7.1.4 证明了当再增加一个单项预测方法时, 基于相关系数的组合预测模型的最大组合预测相关系数可能不变. 这表明组合预测模型可能存在冗余预测方法. 下面定理为冗余信息提供了判定.

定理 7.1.5　在基于相关系数的组合预测模型中, 若第 i 种单项预测方法优超第 k 种单项预测方法, 则第 k 种单项预测方法一定为冗余预测方法, 即组合预测模型的冗余度至少为 $1/m$.

证明　采用反证法. 假设 $\boldsymbol{L}^*=(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型的最优解且第 k 种单项预测方法不为冗余预测方法, 即 $l_k^*>0$, 且 $\sum\limits_{i=1}^{m}l_i^*=1$, 则与 $\boldsymbol{L}^*$ 对应的组合预测模型的目标函数值为

$$R(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)=\frac{\sum\limits_{i=1}^{m}l_i^*\sum\limits_{t=1}^{N}e_te_{it}}{\sqrt{\sum\limits_{t=1}^{N}e_t^2}\sqrt{\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^{*}}}. \tag{7.1.16}$$

构造另外一个向量

$$\hat{\boldsymbol{L}}=\boldsymbol{L}^*+\tilde{\boldsymbol{L}}=(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)^{\mathrm{T}},$$

其中 $\tilde{\boldsymbol{L}}=(0,\cdots,l_k^*,\cdots,-l_k^*,\cdots,0)^{\mathrm{T}}$, 即 $\tilde{\boldsymbol{L}}$ 第 i 个分量为 l_k^*, 第 k 个分量为 $-l_k^*$, 其余 $m-2$ 个分量全为 0. 显然 $\hat{\boldsymbol{L}}$ 为组合预测模型 (7.1.3) 的一可行解, 则与 $\hat{\boldsymbol{L}}$ 对应组合预测模型的目标函数值为

$$R(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)=\frac{\sum_{i=1}^{m}l_i^*\sum_{t=1}^{N}e_te_{it}+l_k^*\sum_{t=1}^{N}e_te_{it}-l_k^*\sum_{t=1}^{N}e_te_{kt}}{\sqrt{\sum_{t=1}^{N}e_t^2}\sqrt{\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\hat{\boldsymbol{L}}}}.\quad(7.1.17)$$

注意到

$$\begin{aligned}\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\hat{\boldsymbol{L}}&=(\boldsymbol{L}^*+\tilde{\boldsymbol{L}})^{\mathrm{T}}\boldsymbol{E}(\boldsymbol{L}^*+\tilde{\boldsymbol{L}})=\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^*+2\tilde{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^*+\tilde{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\tilde{\boldsymbol{L}}\\&=\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^*+2l_k^*\sum_{j\neq k}l_j^*(E_{ij}-E_{kj})+l_k^{*2}(E_{ii}-E_{kk}),\end{aligned}\quad(7.1.18)$$

又因为第 i 种单项预测方法优超第 k 种单项预测方法, 所以由定义 7.1.4 知

$$E_{ii}<E_{kk},\quad E_{ij}<E_{kj},\quad j=1,\cdots,m,\quad j\neq k.$$

由假设 $l_k^*>0$ 和式 (7.1.18) 知

$$\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\hat{\boldsymbol{L}}<\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^*,\quad \sum_{i=1}^{m}l_i^*\sum_{t=1}^{N}e_te_{it}+l_k^*\sum_{t=1}^{N}e_te_{it}-l_k^*\sum_{t=1}^{N}e_te_{kt}>\sum_{i=1}^{m}l_i^*\sum_{t=1}^{N}e_te_{it}.\quad(7.1.19)$$

所以由式 (7.1.16)、(7.1.17)、(7.1.19) 得

$$R(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)>R(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*).$$

而这与 $(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型 (7.1.3) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 此即组合预测模型 (7.1.3) 的冗余度至少为 $1/m$. 证毕.

7.1.4 实例分析[90]

以地方税收收入预测为例说明进行基于相关系数的组合预测模型的有效性. 地方税收收入与许多国民经济指标都有着密切的关系. 这里收集了 1994~2004 年某省三个指标的历史统计资料[89], 包括国内生产总值、社会商品零售总额、地方税收收入. 具体指标数值见表 7.1.1.

表 7.1.1 国内生产总值、社会商品零售总额和地方税收收入指标值及三种单项模型预测值 (单位：亿元)

年份	t	GDP_t	RSV_t	x_t	模型(1)x_{1t}	模型(2)x_{2t}	模型(3)x_{3t}
1994	2	1488.5	453.2	31.93	27.64	37.08	35.79
1995	2	2003.60	586.50	48.07	52.90	55.62	49.76
1996	3	2339.25	727.10	66.32	70.24	67.71	64.50
1997	4	2669.95	859.80	82.58	81.51	79.62	78.41
1998	5	2805.45	924.80	90.49	88.60	84.50	85.23
1999	6	2908.59	979.10	96.30	93.37	88.21	90.92
2000	7	3038.24	1054.30	89.74	97.68	92.88	98.80
2001	8	3290.13	1142.80	108.06	103.42	101.95	108.08
2002	9	3553.56	1228.70	106.93	112.44	111.43	117.09
2003	10	3972.38	1331.20	120.78	126.63	126.51	127.83
2004	11	4812.7	1503.1	152.77	147.83	156.77	145.85

表 7.1.1 中, GDP_t, RSV_t, x_t 分别表示全省各年的国内生产总值、社会商品零售总额、地方税收收入. x_{1t}, x_{2t}, x_{3t} 分别表示三种单项预测模型对地方税收收入预测值, t 表示时间.

本节分别建立地方税收收入与时间、国内生产总值、社会商品零售总额共三种指标的单项预测模型, 利用 SPSS 软件估计出回归模型的参数. 这三种单项预测模型均通过 F 检验, 它们的方程分别是

(1) 地方税收收入与时间的回归预测方程, 记为模型 (1).

$$x_{1t} = -7.422\ 765 + 40.58\ 447t - 5.833\ 866t^2 + 0.311\ 590t^3, \quad t = 1, 2, \cdots, 11.$$

(2) 地方税收收入与国内生产总值的回归预测方程, 记为模型 (2).

$$x_{2t} = -16.517\ 498\ 306\ 4 + 0.036\ 005\ 999\ 585\ 51\ \text{GDP}_t.$$

(3) 地方税收收入与社会商品零售总额的回归预测方程, 记为模型 (3).

$$x_{3t} = -11.719\ 811\ 151\ 93 + 0.104\ 831\ 308\ 9792\ \text{RSV}_t.$$

根据以上三种单项预测模型可分别计算出地方税收收入的预测值, 结果见表 7.1.1 所示.

将表 7.1.1 中数据代入基于相关系数的组合预测模型 (7.1.3) 中, 经计算得如下最优组合预测模型

$$\max R(l_1, l_2, l_3) = \frac{11102l_1 + 10994l_2 + 10977l_3}{\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}}},$$

$$\text{s.t.} \begin{cases} l_1 + l_2 + l_3 = 1, \\ l_1 \geqslant 0, l_2 \geqslant 0, l_3 \geqslant 0, \end{cases}$$

其中组合预测协方差信息矩阵 $\boldsymbol{E}=\begin{bmatrix}11156 & 10954 & 11005\\ 10954 & 11008 & 10884\\ 11005 & 10884 & 11041\end{bmatrix}$, $\boldsymbol{L}=(l_1,l_2,l_3)^{\mathrm{T}}$ 为以上三种单项预测模型在组合预测中的权向量.

利用 Matlab 最优化工具箱计算基于相关系数的最优组合预测模型得组合预测权系数为

$$l_1^*=0.5850,\quad l_2^*=0.4150,\quad l_3^*=0.$$

由此可见单项预测模型 (3) 为冗余预测方法.

如果以常用的基于误差平方和为准则的组合预测模型[15], 可得如下最优化模型

$$\min S(l_1,l_2,l_3)=243.991l_1^2+311.98l_2^2+377.94l_3^2+299.92l_1l_2+435.83l_1l_3+408.64l_2l_3,$$
$$\text{s.t.}\begin{cases}l_1+l_2+l_3=1,\\ l_1\geqslant 0,l_2\geqslant 0,l_3\geqslant 0.\end{cases}$$

利用 Matlab 最优化工具箱计算出相应的最优组合预测权系数为

$$l_1^{**}=0.6328,\quad l_2^{**}=0.3672,\quad l_3^{**}=0.$$

为了反映所提出的组合预测模型的有效性, 选择下列以下两个误差指标评价预测效果.

误差平方和

$$\mathrm{SSE}=\sum_{t=1}^{N}(x_t-\hat{x}_t)^2.$$

平均绝对百分比误差

$$\mathrm{MAPE}=\frac{1}{N}\sum_{t=1}^{N}|(x_t-\hat{x}_t)/x_t|,$$

其中 x_t 为指标序列第 t 时刻的实际观察值, $\hat{x}_t$ 为某种预测方法第 t 时刻的预测值, N 表示样本数量, $t=1,2,\cdots,N$, 按上述两个误差指标分别计算三种单项预测模型的预测误差和基于相关系数的组合预测模型的预测误差, 结果见表 7.1.2.

表 7.1.2 三种单项预测和两种组合预测模型的预测效果的评价

预测效果评价指标	SSE	MAPE	相关系数
单项预测模型(1)	243.99	0.0565	0.9894
单项预测模型(2)	311.98	0.0666	0.9864
单项预测模型(3)	377.94	0.0589	0.9834
基于相关系数的组合预测模型	210.04	0.0466	0.9910
基于误差平方和的组合预测模型	209.46	0.0463	0.9909

从表 7.1.2 可以看出, 基于相关系数的组合预测的两个误差指标值均低于三种单项预测模型预测误差指标值, 同时, 地方税收收入的组合预测值序列与实际值序列的相关系数均高于三种单项预测模型的相关系数, 从而表明本节提出的组合预测方法能够有效地提高预测精度.

另外, 从本例计算结果来看, 本节提出的基于相关系数的组合预测方法和常用的基于误差平方和的组合预测方法在预测精度上相差无几, 因而基于相关系数的组合预测方法不失为另一种有效的组合预测方法.

7.2 基于灰色关联度的组合预测模型的性质

本节主要针对文献 [46] 提出的基于灰色关联度最大化组合预测模型, 提出新的优性组合预测、预测方法优超、冗余度等概念, 研究它的性质, 从而在理论上说明基于灰色关联度的组合预测方法的有效性.

7.2.1 几个概念及基于灰色关联度最大化组合预测模型[91]

设某社会经济现象的指标序列的实际值为 $\{x_t, t=1,2,\cdots,N\}$, 有 m 个单项预测方法对其进行预测, x_{it} 为第 i 种预测方法第 t 时刻的预测值, i=1,2,$\cdots$, m, t=1,2,$\cdots$, N. 则有如下几个概念.

定义 7.2.1[51] 令

$$\gamma_{0i}=\frac{1}{N}\sum_{t=1}^{N}\frac{\min\limits_{1\leqslant i\leqslant m}\min\limits_{1\leqslant t\leqslant N}|e_{it}|+\rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}{|e_{it}|+\rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}.$$

则称 γ_{0i} 为第 i 种单项预测方法预测值序列 $\{x_{it}, t=1,2,\cdots,N\}$ 与实际值序列 $\{x_t, t=1,2,\cdots,N\}$ 的灰色关联度, 简称为第 i 种单项预测方法的灰色关联度, i=1,2,$\cdots$, m. 其中 $\rho\in(0,1)$ 为分辨系数, 通常可取 ρ=0 .5, $e_{it}=x_t-x_{it}$ 为第 i 种预测方法在第 t 时刻的预测误差.

显然, $0\leqslant\gamma_{0i}\leqslant 1$. 根据定义 7.2.1, 只有当某种单项预测方法预测值与实际值序列完全相同时, 即预测百分之百准确, 才有它们之间的灰关联度等于 1.

对某社会经济现象的指标序列进行组合预测, 设 x_t 的组合预测值为

$$\hat{x}_t=l_1x_{1t}+l_2x_{2t}+\cdots+l_mx_{mt},\quad t=1,2,\cdots,N, \tag{7.2.1}$$

其中 $l_1,l_2,\cdots,l_m$ 为 m 种单项预测方法的加权系数, 且满足 $\sum\limits_{i=1}^{m}l_i=1$, $l_i\geqslant 0$, $i=1,2,\cdots,m$. 由式 (7.2.1) 知成立下列不等式

$$\min_{1\leqslant i\leqslant m}x_{it}\leqslant\hat{x}_t\leqslant\max_{1\leqslant i\leqslant m}x_{it},\quad t=1,2,\cdots,N. \tag{7.2.2}$$

设 e_t 为组合预测方法第 t 时刻的预测误差, 注意到 $\sum_{i=1}^{m} l_i = 1$, 则有

$$e_t = x_t - \hat{x}_t = x_t - \sum_{i=1}^{m} l_i x_{it} = \sum_{i=1}^{m} l_i (x_t - x_{it}) = \sum_{i=1}^{m} l_i e_{it}, \quad t = 1, 2, \cdots, N. \tag{7.2.3}$$

设 γ 为组合预测值方法的灰色关联度, 由定义 7.2.1 与式 (7.2.2)、(7.2.3) 得

$$\gamma = \frac{1}{N} \sum_{t=1}^{N} \frac{\min\limits_{1\leqslant i\leqslant m} \min\limits_{1\leqslant t\leqslant N} |e_{it}| + \rho \max\limits_{1\leqslant i\leqslant m} \max\limits_{1\leqslant t\leqslant N} |e_{it}|}{\left|\sum\limits_{i=1}^{m} l_i e_{it}\right| + \rho \max\limits_{1\leqslant i\leqslant m} \max\limits_{1\leqslant t\leqslant N} |e_{it}|}. \tag{7.2.4}$$

显然, 灰色关联度 γ 为各种预测方法的加权系数向量 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 的函数, 所以 γ 可记为 $\gamma(\boldsymbol{L})$.

根据灰色关联度的原理, 组合预测方法的灰色关联度 γ 越大表示组合预测方法越有效, 因此, 基于灰色关联度的组合预测模型可表示成如下模型[46]

$$\max \gamma(\boldsymbol{L}) = \frac{1}{N} \sum_{t=1}^{N} \frac{\min\limits_{1\leqslant i\leqslant m} \min\limits_{1\leqslant t\leqslant N} |e_{it}| + \rho \max\limits_{1\leqslant i\leqslant m} \max\limits_{1\leqslant t\leqslant N} |e_{it}|}{\left|\sum\limits_{i=1}^{m} l_i e_{it}\right| + \rho \max\limits_{1\leqslant i\leqslant m} \max\limits_{1\leqslant t\leqslant N} |e_{it}|},$$

$$\text{s.t.} \begin{cases} \sum\limits_{i=1}^{m} l_i = 1, \\ l_i \geqslant 0, \quad i = 1, 2, \cdots, m. \end{cases} \tag{7.2.5}$$

记 $\gamma_{\min} = \min\limits_{1\leqslant i\leqslant m} \{\gamma_{0i}\}$, $\gamma_{\max} = \max\limits_{1\leqslant i\leqslant m} \{\gamma_{0i}\}$, 即 $\gamma_{\min}$ 表示 m 种单项预测方法的灰色关联度的最小者, $\gamma_{\max}$ 表示其中的最大者.

定义 7.2.2 若 $\gamma < \gamma_{\min}$, 则称组合预测模型 (7.2.5) 为劣性组合预测, 若 $\gamma_{\min} \leqslant \gamma \leqslant \gamma_{\max}$, 则称组合预测模型 (7.2.5) 为非劣性组合预测, 若 $\gamma > \gamma_{\max}$, 则称组合预测模型 (7.2.5) 为优性组合预测.

定义 7.2.2 表明, 只有组合预测灰色关联度大于各单项预测灰色关联度中最大者, 该组合预测模型才为优性组合预测. 优性组合预测就是从灰色关联度这个角度考虑, 它优于各单项预测方法中的最 "好" 者.

在文献 [46] 的实证分析中, 例 1 有两种单项预测方法参与组合预测, 基于灰色关联度的算术平均组合预测模型的最优解为 $l_1^* = 0.2947$, $l_2^* = 0.7053$, 且计算出

$$\gamma_{01} = 0.6323, \quad \gamma_{02} = 0.6813, \quad \gamma(l_1^*, l_2^*) = 0.7496,$$

显然 $\gamma(l_1^*, l_2^*) > \max\{\gamma_{01}, \gamma_{02}\}$, 这表明该最优解对应的组合预测是优性组合预测.

定义 7.2.3 若第 j 种和第 k 种单项预测方法预测误差满足不等式

$$|e_{jt}| \leqslant |e_{kt}|, \quad t = 1, 2, \cdots, N.$$

则称第 j 种单项预测方法优超第 k 种单项预测方法. 若对任意时刻均有严格的不等号成立, 则称第 j 种单项预测方法严格优超第 k 种单项预测方法.

实际上根据定义 7.2.1 知, 第 j 种单项预测方法严格优超第 k 种单项预测方法已蕴涵 $\gamma_{0j} > \gamma_{0k}$. 即就灰色关联度这个指标而言, 直观上可认为第 j 种单项预测方法要 "好于" 第 k 种单项预测方法.

7.2.2 非劣性组合预测和优性组合预测存在的条件[91]

定理 7.2.1 若某种单项预测方法在每个时刻的预测误差绝对值均不小于其他单项预测方法, 则组合预测模型 (7.2.5) 的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 不妨设第一种单项预测方法在每个时刻的预测误差绝对值均不小于其他单项预测方法, 即 $|e_{it}| \leqslant |e_{1t}|, t = 1, 2, \cdots, N, i = 2, 3, \cdots, m$. 则有

$$\frac{\min\limits_{1\leqslant i\leqslant m}\min\limits_{1\leqslant t\leqslant N}|e_{it}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}{|e_{it}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|} \geqslant \frac{\min\limits_{1\leqslant i\leqslant m}\min\limits_{1\leqslant t\leqslant N}|e_{it}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}{|e_{1t}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|},$$

$$t = 1, 2, \cdots, N,$$

所以由定义 7.2.1 知 $\gamma_{0i} \geqslant \gamma_{01}, i = 2, 3, \cdots, m$ 即

$$\gamma_{\min} = \gamma_{01}.$$

设 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为组合预测模型 (7.2.5) 的任一个可行解, 则 $\sum\limits_{i=1}^{m} l_i = 1, l_i \geqslant 0$, $i = 1, 2, \cdots, m$. 所以

$$\left|\sum_{i=1}^{m} l_i e_{it}\right| \leqslant \sum_{i=1}^{m} l_i |e_{it}| \leqslant |e_{1t}| \sum_{i=1}^{m} l_i = |e_{1t}|, \quad t = 1, 2, \cdots, N.$$

由式 (7.2.4) 知 $\gamma \geqslant \gamma_{01}$, 即

$$\gamma \geqslant \gamma_{\min}.$$

再由定义 7.2.2 得结论成立. 证毕.

定理 7.2.1 表明, 从组合预测灰色关联度的角度而言, 在一定的条件下组合预测模型 (7.2.5) 的任一个归一化非负权系数所对应的组合预测均不会比 "最差" 的单项预测方法还要 "差".

定理 7.2.2 设 $\gamma_{01}=\max\limits_{1\leqslant i\leqslant m}\{\gamma_{0i}\}$, 即 γ_{01} 表示 m 种单项预测方法中的最大灰色关联度, $d_t\in(-|e_{1t}|,|e_{1t}|)$, $t=1,2,\cdots,N$, 若如下 $(N+1)\times m$ 的线性方程组

$$\begin{cases} e_{11}l_1+e_{21}l_2+\cdots+e_{m1}l_m=d_1,\\ e_{12}l_1+e_{22}l_2+\cdots+e_{m2}l_m=d_2,\\ \qquad\cdots\cdots\cdots\cdots\\ e_{1N}l_1+e_{2N}l_2+\cdots+e_{mN}l_m=d_N,\\ l_1+l_2+\cdots+l_m=1 \end{cases} \tag{7.2.6}$$

存在非负解, 则组合预测模型 (7.2.5) 一定存在优性组合预测.

证明 设如上的 $(N+1)\times m$ 的线性方程组存在的非负解为 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$, 由 (7.2.6) 式知

$$\sum_{i=1}^m l_i^* e_{it}=d_t,\quad t=1,2,\cdots,N,\quad \sum_{i=1}^m l_i^*=1,\quad l_i^*\geqslant 0.$$

因为 $d_t\in(-|e_{1t}|,|e_{1t}|)$, 所以 $\left|\sum\limits_{i=1}^m l_i^* e_{it}\right|<|e_{1t}|, t=1,2,\cdots,N$. 由 (7.2.4) 式知 $\gamma>\gamma_{01}$, 而 $\gamma_{01}=\max\limits_{1\leqslant i\leqslant m}\{\gamma_{0i}\}$, 即

$$\gamma>\gamma_{\max}.$$

由定义 7.2.2 得结论成立. 证毕.

推论 7.2.1 设 $\gamma_{01}=\max\limits_{1\leqslant i\leqslant m}\{\gamma_{0i}\}$, 若 $\bar{e}_t=\dfrac{1}{m}\sum\limits_{i=1}^m e_{it}\in(-|e_{1t}|,|e_{1t}|)$, $t=1,2,\cdots,N$, 则简单等权平均组合预测方法是优性组合预测方法.

7.2.3 冗余预测方法的一个判定[91]

一般认为随着参与组合预测的单项预测方法个数 m 的增加, 组合预测模型 (7.2.5) 的最大组合预测灰色关联度一定是 m 的严格单调增加函数, 然而实际上, 当再增加一个单项预测方法时, 模型的最大组合预测灰色关联度可能不变. 这是由于在一定的条件下组合预测模型 (7.2.5) 可能存在冗余预测方法.

定理 7.2.3 在组合预测模型 (7.2.5) 中, 若任意第 t 时刻各单项预测方法的预测误差 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号相同, $t=1,2,\cdots,N$, 且第 j 种单项预测方法严格优超第 k 种单项预测方法, 则组合预测模型 (7.2.5) 的冗余度至少为 $1/m$.

证明 采用反证法. 假设 $\boldsymbol{L}^*=(l_1^*,\cdots,l_j^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型 (7.2.5) 的最优解且第 k 种单项预测方法不为冗余预测方法, 即 $l_k^*>0$, $\sum\limits_{i=1}^m l_i^*=1$, 因

为任意第 t 时刻各单项预测方法的预测误差 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的符号相同, 所以有

$$\left|\sum_{i=1}^{m} l_i^* e_{it}\right| = \sum_{i=1}^{m} l_i^* |e_{it}|, \quad t = 1, 2, \cdots, N. \tag{7.2.7}$$

由式 (7.2.5)、(7.2.7) 知, 与 $\boldsymbol{L}^*$ 对应组合预测模型 (7.2.5) 的最优目标函数值为

$$\gamma(\boldsymbol{L}^*) = \frac{1}{N}\sum_{t=1}^{N} \frac{\min\limits_{1\leqslant i\leqslant m}\min\limits_{1\leqslant t\leqslant N}|e_{it}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}{\sum\limits_{i=1}^{m} l_i^*|e_{it}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}. \tag{7.2.8}$$

构造向量

$$\hat{\boldsymbol{L}} = (l_1^*, \cdots, l_{j-1}^*, l_j^* + l_k^*, l_{j+1}^*, \cdots, l_{k-1}^*, 0, l_{k+1}^*, \cdots, l_m^*)^{\mathrm{T}},$$

即 $\hat{\boldsymbol{L}}$ 的第 j 个分量为 $l_j^* + l_k^*$, 第 k 个分量为 0, 其余 $m-2$ 个分量与 $\boldsymbol{L}^* = (l_1^*, \cdots, l_j^*, \cdots, l_k^*, \cdots, l_m^*)^{\mathrm{T}}$ 相同. 显然 $\hat{\boldsymbol{L}}$ 为组合预测模型 (7.2.5) 的一可行解, 则与 $\hat{\boldsymbol{L}}$ 对应模型的目标函数值为

$$\gamma(\hat{\boldsymbol{L}}) = \frac{1}{N}\sum_{t=1}^{N} \frac{\min\limits_{1\leqslant i\leqslant m}\min\limits_{1\leqslant t\leqslant N}|e_{it}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}{\sum\limits_{i\neq j,k} l_i^*|e_{it}| + \left(l_j^* + l_k^*\right)|e_{jt}| + \rho\max\limits_{1\leqslant i\leqslant m}\max\limits_{1\leqslant t\leqslant N}|e_{it}|}. \tag{7.2.9}$$

又因为第 j 种单项预测方法严格优超第 k 种预测方法, 所以由定义 7.2.3 知

$$|e_{jt}| < |e_{kt}|, \quad t = 1, 2, \cdots, N.$$

注意到假设 $l_k^* > 0$, 则有

$$\sum_{i=1}^{m} l_i^*|e_{it}| - \left(\sum_{i\neq j,k} l_i^*|e_{it}| + \left(l_j^* + l_k^*\right)|e_{jt}|\right) = l_k^*\left(|e_{kt}| - |e_{jt}|\right) > 0, \tag{7.2.10}$$

所以

$$\sum_{i=1}^{m} l_i^*|e_{it}| > \sum_{i\neq j,k} l_i^*|e_{it}| + \left(l_j^* + l_k^*\right)|e_{jt}|. \tag{7.2.11}$$

由式 (7.2.8)、(7.2.9)、(7.2.11) 得

$$\gamma(\boldsymbol{L}^*) < \gamma(\hat{\boldsymbol{L}}).$$

这与 $\boldsymbol{L}^* = (l_1^*, \cdots, l_i^*, \cdots, l_k^*, \cdots, l_m^*)^{\mathrm{T}}$ 为组合预测模型 (7.2.5) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 即组合预测模型 (7.2.5) 的冗余度至少为 $1/m$. 证毕.

7.3 基于向量夹角余弦的组合预测模型的性质

本节主要针对基于向量夹角的余弦的组合预测模型, 给出新的优性组合预测、预测方法优超、冗余度等概念, 研究它的性质, 从而在理论上说明基于向量夹角的余弦的组合预测方法的有效性.

7.3.1 符号说明及概念[92]

设某社会经济现象的指标序列的实际值为 $\{x_t, t=1,2,\cdots,N\}$, 设有 m 个单项预测方法对其进行预测, 设第 i 种单项预测方法第 t 时刻的预测值为 x_{it}, i=1,2,$\cdots$, m, t=1,2,$\cdots$, N. 设 $l_1,l_2,\cdots,l_m$ 为 m 种单项预测方法在组合预测中的加权系数, 且满足 $\sum\limits_{i=1}^{m} l_i=1, l_i \geqslant 0, i=1,2,\cdots,m$, 根据加权算术平均的组合预测原理, 则有

$$\hat{x}_t=\sum_{i=1}^{m} l_i x_{it}, \quad t=1,2,\cdots,N, \tag{7.3.1}$$

其中 $\hat{x}_t$ 表示第 t 时刻预测对象实际值 x_t 的组合预测值.

令$\boldsymbol{X}=[x_1,x_2,\cdots,x_N]^{\mathrm{T}}$, $\boldsymbol{X}_i=[x_{i1},x_{i2},\cdots,x_{iN}]^{\mathrm{T}}$, i=1,2,$\cdots$, m, $\hat{\boldsymbol{X}}=[\hat{x}_1,\hat{x}_2,\cdots,\hat{x}_N]^{\mathrm{T}}$, 则$\boldsymbol{X}$表示预测对象实际值向量, $\boldsymbol{X}_i$ 表示第 i 种单项预测方法预测值向量, $\hat{\boldsymbol{X}}$ 表示组合预测值向量.

令

$$\eta_i=\frac{\sum\limits_{t=1}^{N} x_t x_{it}}{\sqrt{\sum\limits_{t=1}^{N} x_t^2}\sqrt{\sum\limits_{t=1}^{N} x_{it}^2}}, \quad i=1,2,\cdots,m, \quad \eta=\frac{\sum\limits_{t=1}^{N} x_t \hat{x}_t}{\sqrt{\sum\limits_{t=1}^{N} x_t^2}\sqrt{\sum\limits_{t=1}^{N} \hat{x}_t^2}}. \tag{7.3.2}$$

根据两个向量夹角的余弦计算公式知, η_i 为第 i 种单项预测方法预测值向量$\boldsymbol{X}_i$与预测对象实际值向量$\boldsymbol{X}$的夹角的余弦, η 为组合预测值向量 $\hat{\boldsymbol{X}}$ 与实际值向量$\boldsymbol{X}$的夹角的余弦.

再令

$$\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}, \quad F_{ij}=\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{X}_j=\sum_{t=1}^{N} x_{it}x_{jt}, \quad i,j=1,2,\cdots,m,$$
$$\boldsymbol{F}=(F_{ij})_{m\times m}, \tag{7.3.3}$$

其中$\boldsymbol{L}$表示组合预测加权系数列向量, F_{ij} 表示向量 $\boldsymbol{X}_i$ 与 $\boldsymbol{X}_j$ 的内积, $\boldsymbol{F}$ 表示 $m\times m$

的方阵, $\boldsymbol{F}$ 称为组合预测信息矩阵. 则有

$$\sum_{t=1}^{N}\left(\sum_{i=1}^{m}l_i x_{it}\right)^2=\sum_{t=1}^{N}\left(\sum_{i=1}^{m}\sum_{j=1}^{m}l_i l_j x_{it}x_{jt}\right)=\sum_{i=1}^{m}\sum_{j=1}^{m}l_i l_j\left(\sum_{t=1}^{N}x_{it}x_{jt}\right)$$
$$=\sum_{i=1}^{m}\sum_{j=1}^{m}l_i l_j F_{ij}=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}. \tag{7.3.4}$$

把式 (7.3.1) 代入到式 (7.3.2) 的 η 中并注意式 (7.3.3)、(7.3.4), 则向量夹角的余弦 η_i, η 可以写成如下形式

$$\eta_i=\frac{\sum_{t=1}^{N}x_t x_{it}}{\sqrt{\sum_{t=1}^{N}x_t^2}\sqrt{F_{ii}}},\quad i=1,2,\cdots,m,\quad \eta=\frac{\sum_{i=1}^{m}l_i\sum_{t=1}^{N}x_t x_{it}}{\sqrt{\sum_{t=1}^{N}x_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}}}, \tag{7.3.5}$$

显然, 组合预测值与实际值向量的向量夹角的余弦 η 为各种单项预测方法的加权系数 $l_1,l_2,\cdots,l_m$ 的函数, 记为 $\eta(l_1,l_2,\cdots,l_m)$. 为了使组合预测值逼近实际值, 我们希望$\boldsymbol{X}$和 $\hat{\boldsymbol{X}}$ 这两个向量之间的夹角愈小愈好. 也就是向量夹角的余弦愈大愈好, 所以 $\eta(l_1,l_2,\cdots,l_m)$ 越大表示组合预测方法越有效. 当 $\hat{\boldsymbol{X}}=\boldsymbol{X}$时, 其向量夹角的余弦达到了最大值 1. 然而组合预测误差也是不可避免的. 因此, 基于向量夹角余弦的最优组合预测模型可表示成如下形式[46]

$$\max\eta(l_1,l_2,\cdots,l_m)=\frac{\sum_{i=1}^{m}l_i\sum_{t=1}^{N}x_t x_{it}}{\sqrt{\sum_{t=1}^{N}x_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}}},$$
$$\text{s.t.}\begin{cases}\sum_{i=1}^{m}l_i=1,\\ l_i\geqslant 0, i=1,2,\cdots,m.\end{cases} \tag{7.3.6}$$

记 $\eta_{\min}=\min\{\eta_i, i=1,2,\cdots,m\}$, $\eta_{\max}=\max\{\eta_i, i=1,2,\cdots,m\}$, 即 $\eta_{\min}$ 表示 m 种单项预测值向量与实际值向量的夹角的余弦中的最小者, $\eta_{\max}$ 表示 m 个向量夹角的余弦中的最大者, 则有如下定义.

定义 7.3.1　设 $\eta(l_1,l_2,\cdots,l_m)$ 为组合预测值向量与实际值向量的夹角的余弦, 若 $\eta(l_1,l_2,\cdots,l_m)<\eta_{\min}$, 则称权系数 $l_1,l_2,\cdots,l_m$ 确定的组合预测模型为劣性组合预测, 若 $\eta_{\min}\leqslant\eta(l_1,l_2,\cdots,l_m)\leqslant\eta_{\max}$, 则称之为非劣性组合预测, 若 $\eta(l_1,l_2,\cdots,l_m)>\eta_{\max}$, 则称之为优性组合预测.

定义 7.3.1 表明, 只有组合预测值向量与预测对象实际值向量的夹角的余弦大于各单项预测值向量与预测对象实际值向量的夹角的余弦中最大者, 该组合预测模型才是优性的.

定义 7.3.2 若第 i 种、第 k 种单项预测方法预测值向量与其他单项预测方法预测值向量以及实际值向量的内积满足如下关系式

$$F_{ij} < F_{kj}, \quad j = 1, 2, \cdots, m, \quad \sum_{t=1}^{N} x_t x_{it} > \sum_{t=1}^{N} x_t x_{kt}.$$

则称基于向量夹角的余弦意义下第 i 种单项预测方法优超第 k 种单项预测方法.

实际上, 由定义 7.3.2 知, $F_{ii} < F_{ki}$, $F_{ik} < F_{kk}$, 由式 (7.3.3) 知 $F_{ki} = F_{ik}$, 则有 $F_{ii} < F_{kk}$. 因此, 若第 i 种单项预测方法优超第 k 种单项预测方法, 则由向量夹角的余弦计算式知 $\eta_i > \eta_k$. 因此就向量夹角的余弦而言, 直观上可以认为第 i 种单项预测方法获得预测值更 "接近" 于实际值, 即第 i 种单项预测方法要 "好于" 第 k 种单项预测方法.

定义 7.3.3 若某种单项预测方法在组合预测模型最优权系数中为零, 则称该单项预测方法为冗余预测方法. 即该种单项预测方法增加到组合预测模型中不能增加组合预测向量夹角的余弦, 表明该种单项预测方法只提供冗余信息.

7.3.2 基于向量夹角的余弦的非劣性组合预测和优性组合预测的存在性[92]

引理 7.3.1 假定 $m(m < N)$ 种单项预测方法的预测值向量组 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 是线性无关的, 则组合预测信息矩阵 $\boldsymbol{F}$ 为正定矩阵.

证明与定理 4.2.1 类似, 故省略. 引理 7.3.1 表明在预测值向量组 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 是线性无关的条件下, 组合预测信息矩阵 $\boldsymbol{F}$ 为正定阵.

定理 7.3.1 假定 $m(m < N)$ 种单项预测方法的预测值向量组 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 是线性无关的, 基于向量夹角的余弦的组合预测模型 (7.3.6) 的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为组合预测模型 (7.3.6) 的任一个可行解, 则有

$$\sum_{i=1}^{m} l_i = 1, \quad l_i \geqslant 0, \quad i = 1, 2, \cdots, m. \tag{7.3.7}$$

设 m 种单项预测方法预测值向量与预测对象实际值向量的向量夹角的余弦按大小排序为 $\eta_1 \geqslant \eta_2 \geqslant \cdots \geqslant \eta_m$, 即

$$\eta_1 = \max\{\eta_i, i = 1, 2, \cdots, m\}, \quad \eta_m = \min\{\eta_i, i = 1, 2, \cdots, m\},$$

所以对任意的 η_i, 均有下式成立 $\eta_i \geqslant \eta_m$, $i=1,2,\cdots,m-1$. 再根据 (7.3.5) 式得

$$\sum_{t=1}^{N} x_t x_{it} \geqslant \eta_m \sqrt{\sum_{t=1}^{N} x_t^2}\sqrt{F_{ii}}, \quad i=1,2,\cdots,m-1.$$

从而

$$\eta = \frac{\sum_{i=1}^{m} l_i \sum_{t=1}^{N} x_t x_{it}}{\sqrt{\sum_{t=1}^{N} x_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}}} \geqslant \frac{\eta_m \sum_{i=1}^{m} l_i \sqrt{\sum_{t=1}^{N} x_t^2}\sqrt{F_{ii}}}{\sqrt{\sum_{t=1}^{N} x_t^2}\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}}} = \frac{\eta_m \sum_{i=1}^{m} l_i \sqrt{F_{ii}}}{\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}}}. \tag{7.3.8}$$

因为向量组 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 是线性无关的, 由引理 7.3.1 知组合预测信息矩阵 $\boldsymbol{F}$ 为正定阵, 所以它的任意二阶子式 $\begin{vmatrix} F_{ii} & F_{ij} \\ F_{ij} & F_{jj} \end{vmatrix} > 0$, 则有 $F_{ij}^2 < F_{ii}F_{jj}$, 注意到 $l_i \geqslant 0, i=1,2,\cdots,m$, 从而

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L} = \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j F_{ij} \leqslant \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \sqrt{F_{ii}}\sqrt{F_{jj}} = \left(\sum_{i=1}^{m} l_i \sqrt{F_{ii}}\right)^2. \tag{7.3.9}$$

由式 (7.3.7)~(7.3.9) 得 $\eta \geqslant \eta_m = \min\{\eta_i, i=1,2,\cdots,m\}$, 再由定义 7.3.1 得结论成立. 证毕.

定理 7.3.1 表明, 从组合预测向量夹角的余弦的角度而言, 组合预测模型 (7.3.6) 的任一个归一化非负权系数所对应的组合预测均不会比 “最差” 的单项预测方法还要 “差”.

推论 7.3.1　简单平均组合预测方法至少是非劣性组合预测.

以向量夹角的余弦作为判断准则, 优性组合预测一定比 “最好” 的单项预测方法还要 “好”. 因此研究优性组合预测的存在性具有一定的理论意义.

定理 7.3.2　假定 $m(m<N)$ 种单项预测方法的预测值向量组 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 是线性无关的, 基于向量夹角余弦的组合预测模型的冗余度 $k < (m-1)/m$, 则最优解对应的组合预测一定是优性组合预测.

证明　设 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$ 为模型 (7.3.6) 的最优解, $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为其某一个可行解而不是最优解, 则有

$$\sum_{i=1}^{m} l_i^* = 1, \quad l_i^* \geqslant 0, \quad \sum_{i=1}^{m} l_i = 1, \quad l_i \geqslant 0, \quad i=1,2,\cdots,m. \tag{7.3.10}$$

因为基于向量夹角余弦的组合预测模型 (7.3.6) 目标函数是求最大值的, 则有

$$\eta(l_1^*, l_2^*, \cdots, l_m^*) > \eta(l_1, l_2, \cdots, l_m). \tag{7.3.11}$$

由式 (7.3.9)、(7.3.10) 知

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L} \leqslant \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \sqrt{F_{ii}}\sqrt{F_{jj}} \leqslant \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \frac{F_{ii}+F_{jj}}{2} = \sum_{i=1}^{m} l_i F_{ii}.$$

从而有

$$\eta(l_1, l_2, \cdots, l_m) = \frac{\sum_{i=1}^{m} l_i \sum_{t=1}^{N} x_t x_{it}}{\sqrt{\sum_{t=1}^{N} x_t^2 \sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}}}} \geqslant \frac{\sum_{i=1}^{m} l_i \sum_{t=1}^{N} x_t x_{it}}{\sqrt{\sum_{t=1}^{N} x_t^2}\sqrt{\sum_{i=1}^{m} l_i F_{ii}}}. \tag{7.3.12}$$

因为组合预测模型的冗余度 $k < (m-1)/m$, 所以最优解中至少有两个非 0 分量, 从而 $(1,0,\cdots,0)^{\mathrm{T}}$, $(0,1,\cdots,0)^{\mathrm{T}}$, $\cdots$, $(0,0,\cdots,1)^{\mathrm{T}}$ 这 m 个 m 维单位列向量分别是组合预测模型 (7.3.6) 的可行解而不是最优解, 由式 (7.3.11)、(7.3.12) 知

$$\begin{cases} \eta(l_1^*, l_2^*, \cdots, l_m^*) > \eta(1,0,\cdots,0) \geqslant \eta_1, \\ \eta(l_1^*, l_2^*, \cdots, l_m^*) > \eta(0,1,\cdots,0) \geqslant \eta_2, \\ \qquad\cdots\cdots\cdots\cdots \\ \eta(l_1^*, l_2^*, \cdots, l_m^*) > \eta(0,0,\cdots,1) \geqslant \eta_m. \end{cases}$$

所以 $\eta(l_1^*, l_2^*, \cdots, l_m^*) > \max\{\eta_i, i=1,2,\cdots,m\} = \eta_{\max}$, 再由定义 7.3.1 得结论成立.

定理 7.3.2 表明, 在 m 种单项预测方法参与的基于向量夹角的余弦的组合预测模型中, 最优解里只要有两个以上的分量不为 0, 即两个以上单项预测方法提供有效信息, 则最优解所对应的组合预测方法要 "好于" 单项预测方法中的 "最好" 者, 可见, 最优组合预测模型确实综合利用了单项预测方法提供的信息.

7.3.3 基于向量夹角的余弦的冗余预测方法的存在性及其判定[92]

对于基于向量夹角的余弦的组合预测模型 (7.3.6) 有如下结论成立.

定理 7.3.3 基于向量夹角的余弦的组合预测模型的最优目标函数是参与组合预测的各单项预测方法总个数 m 的单调不减函数, 即

$$\eta(l_1^*, l_2^*, \cdots, l_m^*) \leqslant \eta(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1}),$$

其中 $\eta(l_1^*, l_2^*, \cdots, l_m^*)$ 表示 m 个单项预测方法参与的基于向量夹角的余弦的组合预测模型对应的最大向量夹角的余弦, $\eta(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1})$ 表示再增加一个单项预测方法共 $m+1$ 个单项预测方法对应的最大向量夹角的余弦.

证明 设 $\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$ 为 m 个单项预测方法参与的组合预测模型 (7.3.6) 的最优解, 其最优目标函数值为

$$\eta(l_1^*,l_2^*,\cdots,l_m^*)=\frac{\sum_{i=1}^{m}l_i^*\sum_{t=1}^{N}x_tx_{it}}{\sqrt{\sum_{t=1}^{N}x_t^2}\sqrt{\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{F}\boldsymbol{L}^*}},\tag{7.3.13}$$

其中 $\sum_{i=1}^{m}l_i^*=1,\ l_i^*\geqslant 0,\ i=1,2,\cdots,m.$

设 $\bar{\boldsymbol{L}}=(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})^{\mathrm{T}}$ 为再增加一个单项预测方法, 共 $m+1$ 个单项预测方法参与的基于向量夹角的余弦的组合预测模型的最优解, 其最优目标函数值为

$$\eta(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})=\frac{\sum_{i=1}^{m+1}\bar{l}_i\sum_{t=1}^{N}x_tx_{it}}{\sqrt{\sum_{t=1}^{N}x_t^2}\sqrt{\bar{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{F}_{m+1}\bar{\boldsymbol{L}}}},\tag{7.3.14}$$

其中 $\boldsymbol{F}_{m+1}=\begin{pmatrix}\boldsymbol{F} & \boldsymbol{f}\\ \boldsymbol{f}^{\mathrm{T}} & F_{(m+1)(m+1)}\end{pmatrix}$, $\boldsymbol{f}=\left(F_{1(m+1)},F_{2(m+1)},\cdots,F_{m(m+1)}\right)^{\mathrm{T}}$, $F_{i(m+1)}=\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{X}_{m+1}=\sum_{t=1}^{N}x_{it}x_{(m+1)t}$, $i=1,2,\cdots,m+1$, $F_{i(m+1)}$ 表示第 i 种单项预测方法与第 $m+1$ 种单项预测方法预测值向量的内积, $\boldsymbol{F}_{m+1}$ 表示 $m+1$ 个单项预测方法参与的 $m+1$ 阶组合预测信息矩阵. 且 $\sum_{i=1}^{m+1}\bar{l}_i=1,\bar{l}_i\geqslant 0,\ i=1,2,\cdots,m+1.$

令 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*,0)^{\mathrm{T}}$, 显然 $\boldsymbol{L}$ 为 $m+1$ 个单项预测方法参与的基于向量夹角的余弦的组合预测模型的可行解, 则有

$$\eta(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})\geqslant\eta(l_1^*,l_2^*,\cdots,l_m^*,0),$$

由式 (7.3.14) 得

$$\eta(l_1^*,l_2^*,\cdots,l_m^*,0)=\eta(l_1^*,l_2^*,\cdots,l_m^*).$$

所以 $\eta(l_1^*,l_2^*,\cdots,l_m^*)\leqslant\eta(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})$, 即结论成立.

定理 7.3.3 证明了当再增加一个单项预测方法时, 基于向量夹角的余弦的组合预测模型的最优目标函数值可能不变. 这表明组合预测模型可能存在冗余预测方法. 下面一个定理为冗余信息提供了判定.

定理 7.3.4 在基于向量夹角的余弦的组合预测模型中, 若第 i 种单项预测方法优超第 k 种单项预测方法, 则组合预测模型的冗余度至少为 $1/m$.

证明 反证法. 假设 $\boldsymbol{L}^*=(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型的最优解且第 k 种单项预测方法不为冗余预测方法, 即 $l_k^*>0$, $\sum\limits_{i=1}^m l_i^*=1$, 则与 $\boldsymbol{L}^*$ 对应的组合预测模型的目标函数值为

$$\eta(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)=\frac{\sum\limits_{j=1}^m l_j^*\sum\limits_{t=1}^N x_t x_{jt}}{\sqrt{\sum\limits_{t=1}^N x_t^2}\sqrt{\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{F}\boldsymbol{L}^*}}. \tag{7.3.15}$$

构造另外一个向量 $\hat{\boldsymbol{L}}=\boldsymbol{L}^*+\tilde{\boldsymbol{L}}=(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)$, 其中 $\tilde{\boldsymbol{L}}=(0,\cdots,l_k^*,\cdots,-l_k^*,\cdots,0)$, 即 $\tilde{\boldsymbol{L}}$ 的第 i 个分量为 l_k^*, 第 k 个分量为 $-l_k^*$, 其余 $m-2$ 个分量全为 0.

显然, $\hat{\boldsymbol{L}}$ 为组合预测模型 (7.3.1) 的一可行解, 则与 $\hat{\boldsymbol{L}}$ 对应组合预测模型的目标函数值为

$$\eta(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)=\frac{\sum\limits_{j=1}^m l_j^*\sum\limits_{t=1}^N x_t x_{jt}+l_k^*\sum\limits_{t=1}^N x_t x_{it}-l_k^*\sum\limits_{t=1}^N x_t x_{kt}}{\sqrt{\sum\limits_{t=1}^N x_t^2}\sqrt{\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{F}\hat{\boldsymbol{L}}}}. \tag{7.3.16}$$

而

$$\begin{aligned}\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{F}\hat{\boldsymbol{L}}&=(\boldsymbol{L}^*+\tilde{\boldsymbol{L}})^{\mathrm{T}}\boldsymbol{F}(\boldsymbol{L}^*+\tilde{\boldsymbol{L}})=\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{F}\boldsymbol{L}^*+2\tilde{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}^*+\tilde{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{F}\tilde{\boldsymbol{L}}\\&=\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{F}\boldsymbol{L}^*+2l_k^*\sum_{j\neq k}l_j^*(F_{ij}-F_{kj})+l_k^{*2}(F_{ii}-F_{kk}).\end{aligned} \tag{7.3.17}$$

又因为第 i 种单项预测方法优超第 k 种单项预测方法, 所以由定义 7.3.2 知

$$F_{ii}<F_{kk},\quad F_{ij}<F_{kj},\quad j=1,\cdots,m,\quad j\neq k,\quad \sum_{t=1}^N x_t x_{it}>\sum_{t=1}^N x_t x_{kt}.$$

注意到假设 $l_k^*>0$, 且由式 (7.3.17) 知 $\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{F}\hat{\boldsymbol{L}}<\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{F}\boldsymbol{L}^*$, 且

$$\sum_{j=1}^m l_j^*\sum_{t=1}^N x_t x_{jt}+l_k^*\sum_{t=1}^N x_t x_{it}-l_k^*\sum_{t=1}^N x_t x_{kt}>\sum_{j=1}^m l_j^*\sum_{t=1}^N x_t x_{jt}. \tag{7.3.18}$$

由式 (7.3.15)、(7.3.16)、(7.3.18) 得

$$\eta(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)>\eta(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*).$$

而这与 $(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型 (7.3.1) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 此即组合预测模型 (7.3.1) 的冗余度至少为 $1/m$.

7.3.4 实例分析[92]

太阳活动规律的模拟与预测是全球多学科研究的热点问题之一, 尤其是太阳黑子相对数极大极小的年份和数值更为世界所关注[55]. 文献 [55] 利用太阳黑子数从 1700~1979 年共 280 年的时间序列数据建立了五种时间序列模型 (原始时间序列数据参见文献 [50]), 这里取其中的三种时间序列模型进行基于向量夹角的余弦的组合预测来说明本模型的有效性. 这三种时间序列模型[55] 表示如下.

(1) 非线性门限自回归模型 (TAR). 设太阳黑子数时间序列为$\{Z_t,\ t=1,\ 2,\cdots,\ 280\}$, 对$\{Z_t,\ t=1,\ 2,\cdots,\ 280\}$进行 Box-Cox 变换得$\{X_t,\ t=1,\ 2,\cdots,\ 280\}$, 即令

$$X_t=(Z_t^{0.5}-1)/0.5=2(\sqrt{Z_t}-1),\quad t=1,2,\cdots,280.$$

再对$\{X_t\}$建立 TAR 模型得

$$X_t=\begin{cases}1.9659+0.8052X_{t-1}+0.1053X_{t-2}-0.2681X_{t-3}+0.107X_{t-4}-0.1844X_{t-5}\\ \qquad -0.0199X_{t-6}+0.1772X_{t-7}-0.2258X_{t-8}+0.1190X_{t-9}+0.1409X_{t-10}\\ \qquad +\varepsilon_t^{(1)}, & \text{当}X_{t-8}\leqslant 12,\\ 4.8751+1.4196X_{t-1}-0.8261X_{t-2}+\varepsilon_t^{(2)}, & \text{当}X_{t-8}>12.\end{cases}$$

(2) 不含趋势的叠合模型

$$Z_t=-\mathrm{e}^{0.003824t}\left(9.189\sin\frac{2\pi t}{11}+9.841\cos\frac{2\pi t}{11}\right)+52.26+X_t,$$

其中 $X_t=1.168X_{t-1}-0.4283X_{t-2}-0.1561X_{t-3}+0.1874X_{t-4}-0.1115X_{t-5}+0.03894X_{t-6}+0.01896X_{t-7}-0.0192X_{t-8}+0.1955X_{t-9}+\varepsilon_t$.

(3) 自回归模型 (AR(p)). 令 $X_t=Z_t-48.9886$, 对$\{X_t\}$建立 AR(9) 模型得

$$\begin{aligned}X_t=&1.1764X_{t-1}-0.4331X_{t-2}-0.1640X_{t-3}+0.1827X_{t-4}-0.1278X_{t-5}\\ &+0.0406X_{t-6}+0.0158X_{t-7}-0.0305X_{t-8}+0.2192X_{t-9}+\varepsilon_t.\end{aligned}$$

利用以上三种时间序列模型对太阳黑子数预测值[55]见表 7.3.1.

将表 7.3.1 中数据代入到模型 (7.3.6) 中, 经计算得如下基于向量夹角的余弦的最优组合预测模型

$$\max\eta(l_1,l_2)=\frac{24243.53l_1+249.07l_2+227.57l_3}{\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{L}}},$$

$$\text{s.t.}\begin{cases}l_1+l_2+l_3=1,\\ l_1\geqslant 0,l_2\geqslant 0,l_3\geqslant 0,\end{cases}$$

表 7.3.1 太阳黑子数的观察值和三种时间序列模型的预测值

年份	黑子数的观察值	TAR 模型	不含趋势叠合模型	AR 模型
1980	154.6	158.63	163.5	158.97
1981	140.4	137.24	139.5	128.99
1982	115.9	98.21	100.4	84.74
1983	66.6	61.31	65.38	48.64
1984	45.9	32.64	35.02	21.66
1985	17.9	17.42	15.2	9.38
1986	13.4	16.4	7.812	10.7
1987	29.2	26.47	21.75	32.23

其中组合预测信息矩阵 $\boldsymbol{F}=\begin{bmatrix}59741 & 61061 & 56123\\ 61061 & 62539 & 57359\\ 56123 & 57359 & 53167\end{bmatrix}$, $\boldsymbol{L}=(l_1,l_2,l_3)^{\mathrm{T}}$ 为以上三种时间序列预测模型在组合预测中的权向量.

利用 Matlab 最优化工具箱计算得最优组合预测权系数为

$$l_1^*=0.5003,\quad l_2^*=0.4997,\quad l_3^*=0.$$

由此可见自回归预测模型为冗余预测方法.

为了反映组合预测模型 (7.3.6) 的有效性, 通常选择下列预测效果评价指标体系[46]:

(1) 平方和误差 $\mathrm{SSE}=\sum_{t=1}^{N}(x_t-\hat{x}_t)^2$.

(2) 平均绝对误差 $\mathrm{MAE}=\dfrac{1}{N}\sum_{t=1}^{N}|x_t-\hat{x}_t|$.

(3) 平均绝对百分比误差 $\mathrm{MAPE}=\dfrac{1}{N}\sum_{t=1}^{N}|(x_t-\hat{x}_t)/x_t|$.

根据上面的预测效果评价指标公式, 组合预测模型 (7.3.6) 和三种单项预测模型的误差指标的计算结果如表 7.3.2 所示.

表 7.3.2 组合预测模型和三种单项预测模型预测效果评价指标

	SSE	MAE	MAPE
非线性门限自回归模型	559.6576	6.2050	0.1142
不含趋势的叠合模型	534.1510	6.6423	0.1595
自回归模型	2119.4	12.9238	0.2447
组合预测模型	507.7094	6.0483	0.1088

从表 7.3.2 的预测效果评价指标体系来看, 组合预测模型 (7.3.6) 的误差指标值

均低于三种单项预测模型，从而表明基于向量夹角的余弦的最优组合预测模型可以取得比各单项预测方法更好的预测效果.

7.4 基于 Theil 不等系数的优性组合预测模型的性质研究

本节针对提出的基于改进的 Theil 不等系数的组合预测模型，给出新的优性组合预测、预测方法优超等概念，研究了它的基本性质，从而在理论上说明基于改进的 Theil 不等系数的组合预测方法的有效性.

7.4.1 符号说明及基于改进的 Theil 不等系数的组合预测模型概念[93]

设某社会经济现象的指标序列的实际值为$\{x_t, t=1,2,\cdots,N\}$，设有 m 个单项预测方法对其进行预测，这 m 个单项预测方法均参与到组合预测中. 设第 i 种单项预测方法第 t 时刻的预测值为 $x_{it}, t=1,2,\cdots,N$, $i=1,2,\cdots,m$. 设第 i 种单项预测方法在组合预测中的加权系数为 l_i，且满足

$$\sum_{i=1}^{m} l_i = 1, \quad l_i \geqslant 0, \quad i=1,2,\cdots,m. \tag{7.4.1}$$

根据加权算术平均的组合预测原理，则有

$$\hat{x}_t = \sum_{i=1}^{m} l_i x_{it}, \quad t=1,2,\cdots,N, \tag{7.4.2}$$

其中 $\hat{x}_t$ 表示指标序列第 t 时刻的组合预测值.

定义 7.4.1 令

$$\mu_i = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t - x_{it})^2} \Bigg/ \left(\sqrt{\frac{1}{N}\sum_{t=1}^{N}x_t^2} + \sqrt{\frac{1}{N}\sum_{t=1}^{N}x_{it}^2}\right), \qquad i=1,2,\cdots,m,$$

$$\mu = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t - \hat{x}_t)^2} \Bigg/ \left(\sqrt{\frac{1}{N}\sum_{t=1}^{N}x_t^2} + \sqrt{\frac{1}{N}\sum_{t=1}^{N}\hat{x}_t^2}\right).$$

则称 μ_i 为第 i 种单项预测方法预测值序列$\{x_{it}, t=1,2,\cdots,N\}$与指标序列实际值$\{x_t, t=1,2,\cdots,N\}$的 Theil 不等系数，称 μ 为组合预测值序列与指标序列实际值的 Theil 不等系数.

显然，Theil 不等系数 $\mu_i \in [0,1]$ 及 $\mu \in [0,1]$，μ 值越小表示预测精度越高，特别地，当 $\mu=0$ 时，表示组合预测值与指标序列实际值相等，此为理想的或完美的预测. 当 $\hat{x}_t = -x_t$ 时，即组合预测值与实际值完全相反，此时 $\mu=1$，它表示预测极不

准确. 但是当 $\hat{x}_t$ 恒为零时, 也会出现 $\mu=1$ 的情形. 为了区分这两种不同的情形, 则引进如下改进的 Theil 不等系数.

定义 7.4.2 令

$$\begin{cases} \tau_i = \sqrt{\sum_{t=1}^{N}(x_t - x_{it})^2} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2}, \quad i=1,2,\cdots,m, \\ \tau = \sqrt{\sum_{t=1}^{N}(x_t - \hat{x}_t)^2} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2}. \end{cases} \tag{7.4.3}$$

则称 τ_i 和 τ 分别为第 i 种单项预测方法预测值序列$\{x_{it}, t=1,2,\cdots,N\}$及组合预测值序列$\{\hat{x}_t, t=1,2,\cdots,N\}$与实际值序列$\{x_t, t=1,2,\cdots,N\}$改进的 Theil 不等系数.

显然, 改进的 Theil 不等系数 $\tau_i \in [0,\infty)$ 及 $\tau \in [0,\infty)$, τ 值越小表示组合预测精度越高, 当 $\tau=0$ 时, $\hat{x}_t = x_t$, $t=1,2,\cdots,N$, 表示组合预测准确无误. 当 $\hat{x}_t = -x_t$ 时, $\tau=2$, 当 $\hat{x}_t=0$ 时, $\tau=1$, 所以改进的 Theil 不等系数能将它们区分开.

令$e_{it}=x_t-x_{it}$, $i=1,2,\cdots,m$, $t=1,2,\cdots,N$, $e_t=x_t-\hat{x}_t$, 则e_{it}表示第i种预测方法在第 t 时刻的预测误差, e_t表示组合预测方法在第 t 时刻的预测误差, $t=1,2,\cdots,N$.

再令

$$\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}, \quad \boldsymbol{E}=(E_{ij})_{m\times m}.$$

则 $\boldsymbol{L}$ 表示组合预测加权系数列向量, $\boldsymbol{E}$ 表示 $m\times m$ 的方阵, $\boldsymbol{E}$ 称为组合预测误差信息矩阵. 其中当 $i\neq j$ 时, $E_{ij}=\sum_{t=1}^{N} e_{it}e_{jt}$ 表示第 i 种与第 j 种单项预测方法预测误差序列的协方差, 显然 $E_{ij}=E_{ji}$. 当 $i=j$ 时, $E_{ii}=\sum_{t=1}^{N} e_{it}^2$ 表示第 i 种单项预测方法误差序列的方差.

在上述符号下并注意式 (7.4.1)、(7.4.2), 则有

$$\begin{aligned} \sum_{t=1}^{N}(x_t-\hat{x}_t)^2 &= \sum_{t=1}^{N}\left(\sum_{i=1}^{m} l_i(x_t-x_{it})\right)^2 = \sum_{t=1}^{N}\left(\sum_{i=1}^{m} l_i e_{it}\right)^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \left(\sum_{t=1}^{N} e_{it}e_{jt}\right) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j E_{ij} = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}. \end{aligned}$$

所以改进的 Theil 不等系数式 (7.4.3) 可写成下式

$$\tau_i = \sqrt{E_{ii}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2}, \quad \tau = \sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2}. \tag{7.4.4}$$

显然, 组合预测值序列与实际值序列改进的 Theil 不等系数为组合预测方法的加权系数 $l_1, l_2, \cdots, l_m$ 的函数, 记为 $\tau(l_1, l_2, \cdots, l_m)$. 当从改进的 Theil 不等系数角度考察组合预测问题的时候, 我们希望 $\tau(l_1, l_2, \cdots, l_m)$ 愈小愈好, $\tau(l_1, l_2, \cdots, l_m)$ 越小表示组合预测方法越有效. 当 $\hat{x}_t = x_t$, t=1,2,$\cdots$, N, Theil 不等系数达到了最小值 0. 然而预测误差是不可避免的, 因此, 基于改进的 Theil 不等系数的组合预测模型可表示成如下模型

$$\begin{aligned} &\min \tau(l_1, l_2, \cdots, l_m) = \sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2}, \\ &\text{s.t.} \begin{cases} \sum\limits_{i=1}^{m} l_i = 1, \\ l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \end{aligned} \tag{7.4.5}$$

记 $\tau_{\min} = \min\{\tau_i, i = 1, 2, \cdots, m\}$, $\tau_{\max} = \max\{\tau_i, i = 1, 2, \cdots, m\}$, 即 $\tau_{\min}$ 和 $\tau_{\max}$ 分别表示 m 种单项预测值序列与指标序列实际值的改进的 Theil 不等系数中的最小者和最大者, 有如下定义.

定义 7.4.3　若 $\tau(l_1, l_2, \cdots, l_m) > \tau_{\max}$, 则称权系数 $l_1, l_2, \cdots, l_m$ 确定的组合预测模型为劣性组合预测, 若 $\tau_{\min} \leqslant \tau(l_1, l_2, \cdots, l_m) \leqslant \tau_{\max}$, 则称之为非劣性组合预测, 若 $\tau(l_1, l_2, \cdots, l_m) < \tau_{\min}$, 则称为优性组合预测.

定义 7.4.3 表明, 只有组合预测值序列与指标序列实际值的 Theil 不等系数小于各单项预测值序列与指标序列实际值的 Theil 不等系数中最小者, 该组合预测模型才是优性的.

定义 7.4.4　若第 i 种、第 k 种单项预测方法与其他单项预测方法预测值序列以及实际值序列的协方差满足

$$E_{ij} < E_{kj}, \quad j = 1, 2, \cdots, m. \tag{7.4.6}$$

则称第 i 种单项预测方法基于 Theil 不等系数优超第 k 种单项预测方法.

实际上, 由式 (7.4.6) 知 $E_{ii} < E_{ki}$, $E_{ik} < E_{kk}$, 由 E_{ij} 的定义式知 $E_{ki} = E_{ik}$, 则有 $E_{ii} < E_{kk}$, 因此若第 i 种单项预测方法优超第 k 种单项预测方法, 则由改进的 Theil 不等系数计算式 (7.4.4) 知 $\tau_i < \tau_k$. 因此, 就改进的 Theil 不等系数而言, 可以认为第 i 种单项预测方法的预测值更 “接近” 于实际值, 即第 i 种单项预测方法要 “好于” 第 k 种单项预测方法.

7.4.2 基于改进的 Theil 不等系数的组合预测模型的性质[93]

引理 7.4.1 假定 $m(m<N)$ 种单项预测方法的预测误差向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 其中 $\boldsymbol{e}_i=(e_{i1},e_{i2},\cdots,e_{iN})^{\mathrm{T}}$, $i=1,2,\cdots,m$, 则组合预测误差信息矩阵 $\boldsymbol{E}$ 为正定矩阵.

定理 7.4.1 在引理 7.4.1 的假定条件下, 基于改进的 Theil 不等系数的组合预测模型的任一个可行解对应的组合预测至少是非劣性组合预测.

证明 设 $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为组合预测模型 (7.4.5) 的任一个可行解, 设 m 种单项预测方法预测值序列与指标序列实际值改进的 Theil 不等系数按大小排序为

$$\tau_1\geqslant\tau_2\geqslant\cdots\geqslant\tau_m.$$

即 $\tau_1=\max\{\tau_i,i=1,2,\cdots,m\}$, $\tau_m=\min\{\tau_i,i=1,2,\cdots,m\}$, 所以对任意的 τ_i, 均成立

$$\tau_i\leqslant\tau_1,\quad i=2,3,\cdots,m. \tag{7.4.7}$$

把式 (7.4.4) 代入式 (7.4.7) 得

$$\sqrt{E_{ii}}\leqslant\tau_1\sqrt{\sum_{t=1}^{N}x_t^2},\quad i=1,2,\cdots,m. \tag{7.4.8}$$

因为向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 是线性无关的, 由引理 7.4.1 知组合预测误差信息矩阵 $\boldsymbol{E}$ 为正定阵, 所以它的任意二阶主子式 $\begin{vmatrix}E_{ii}&E_{ij}\\E_{ij}&E_{jj}\end{vmatrix}>0$, 则有 $E_{ij}^2<E_{ii}E_{jj}$, 注意到 $l_i\geqslant0,i=1,2,\cdots,m$, 有

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}=\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_jE_{ij}\leqslant\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_j\sqrt{E_{ii}}\sqrt{E_{jj}}=\left(\sum_{i=1}^{m}l_i\sqrt{E_{ii}}\right)^2. \tag{7.4.9}$$

由 (7.4.1)、(7.4.8)、(7.4.9) 式得

$$\begin{aligned}\tau=\sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}}\bigg/\sqrt{\sum_{t=1}^{N}x_t^2}&\leqslant\sum_{i=1}^{m}l_i\sqrt{E_{ii}}\bigg/\sqrt{\sum_{t=1}^{N}x_t^2}\\&\leqslant\tau_1\sum_{i=1}^{m}l_i\sqrt{\sum_{t=1}^{N}x_t^2}\bigg/\sqrt{\sum_{t=1}^{N}x_t^2}=\tau_1.\end{aligned} \tag{7.4.10}$$

从而 $\tau\leqslant\tau_1=\max\{\tau_i,i=1,2,\cdots,m\}$, 再由定义 7.4.3 得结论成立.

定理 7.4.1 表明, 从组合预测改进的 Theil 不等系数的角度而言, 组合预测模型 (7.4.5) 的任意归一化非负权系数所对应的组合预测均不会比 “最差” 的单项预测方法还要 “差”.

推论 7.4.1　简单算术平均组合预测方法至少是非劣性组合预测.

定理 7.4.2　假定 $m(m<N)$ 种单项预测方法的预测误差向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_m$ 是线性无关的, 其中 $\boldsymbol{e}_i=(e_{i1}, e_{i2}, \cdots, e_{iN})^{\mathrm{T}}$, $i=1,2,\cdots,m$, 基于 Theil 不等系数的组合预测模型的冗余度 $k \neq (m-1)/m$, 则其最优解对应的组合预测一定是优性组合预测.

证明　设基于 Theil 不等系数的组合预测模型 (7.4.5) 的最优解为 $\boldsymbol{L}^*=(l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$, 设 $\boldsymbol{L}=(l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为其一个可行解, 则模型 (7.4.5) 的最优目标函数值 $\tau(l_1^*, l_2^*, \cdots, l_m^*)$ 满足

$$\tau(l_1^*, l_2^*, \cdots, l_m^*) < \tau(l_1, l_2, \cdots, l_m). \tag{7.4.11}$$

由式 (7.4.9) 并注意到 $\sum_{i=1}^{m} l_i=1$, $l_i \geqslant 0$, $i=1,2,\cdots,m$, 则有

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L} \leqslant \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \sqrt{E_{ii}}\sqrt{E_{jj}} \leqslant \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \frac{E_{ii}+E_{jj}}{2} = \sum_{i=1}^{m} l_i E_{ii}.$$

所以

$$\tau(l_1, l_2, \cdots, l_m) = \sqrt{\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2} \leqslant \sqrt{\sum_{i=1}^{m} l_i E_{ii}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2}. \tag{7.4.12}$$

由式 (7.4.12) 及式 (7.4.4) 知

$$\begin{cases} \tau(1,0,\cdots,0) \leqslant \sqrt{E_{11}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2} = \tau_1, \\ \tau(0,1,\cdots,0) \leqslant \sqrt{E_{22}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2} = \tau_2, \\ \qquad \cdots\cdots\cdots\cdots \\ \tau(0,0,\cdots,1) \leqslant \sqrt{E_{mm}} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2} = \tau_m. \end{cases} \tag{7.4.13}$$

因为组合预测模型的冗余度 $k \neq (m-1)/m$, 所以最优解中 $\boldsymbol{L}^*=(l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$ 至少有两个非 0 分量, 而 $\boldsymbol{L}_1=(1,0,\cdots,0)^{\mathrm{T}}$ 只有一个非 0 分量, 所以它是组合预测模型 (7.4.5) 的可行解而不是最优解. 同理 $\boldsymbol{L}_2=(0,1,\cdots,0)^{\mathrm{T}}, \cdots, \boldsymbol{L}_m=(0,0,\cdots,1)^{\mathrm{T}}$

也均是组合预测模型 (7.4.5) 的可行解而不是最优解, 分别把它们代入式 (7.4.11) 得

$$\begin{cases} \tau(l_1^*, l_2^*, \cdots, l_m^*) < \tau(1, 0, \cdots, 0), \\ \tau(l_1^*, l_2^*, \cdots, l_m^*) < \tau(0, 1, \cdots, 0), \\ \cdots\cdots\cdots\cdots \\ \tau(l_1^*, l_2^*, \cdots, l_m^*) < \tau(0, 0, \cdots, 1), \end{cases} \tag{7.4.14}$$

所以式 (7.4.13)、(7.4.14) 知

$$\tau(l_1^*, l_2^*, \cdots, l_m^*) < \min\{\tau_i, i = 1, 2, \cdots, m\} = \tau_{\min}.$$

再由定义 7.4.3 得结论成立.

定理 7.4.2 表明, 在 m 种单项预测方法参与的基于 Theil 不等系数的组合预测模型中, 最优解里只要有两个以上的分量不为 0, 即两个以上单项预测方法提供有效信息, 则最优解所对应的组合预测方法要 "好于" 单项预测方法中的 "最好" 者, 可见, 最优组合预测模型确实综合利用了单项预测方法提供的信息.

在文献 [46] 的实证分析中, 例 1 有两种单项预测方法参与组合预测, 基于 Theil 不等系数的算术平均组合预测模型的最优解为

$$l_1^* = 0.5581, \quad l_2^* = 0.4419.$$

这表明组合预测模型的冗余度 $k = 0$, 且计算出

$$\tau_1 = 0.0325, \quad \tau_2 = 0.0360, \quad \tau(l_1^*, l_2^*) = 0.0260.$$

显然, $\tau(l_1^*, l_2^*) < \min\{\tau_1, \tau_2\}$, 该最优解对应的组合预测一定是优性组合预测. 文献 [46] 的例 2 也有同样的结论. 可见文献 [46] 的实证分析验证了定理 7.4.2 结论的正确性.

定理 7.4.3 基于改进的 Theil 不等系数的组合预测模型的最优目标函数是参与组合预测的各单项预测方法总个数 m 的单调不增函数, 即

$$\tau(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1}) \leqslant \tau(l_1^*, l_2^*, \cdots, l_m^*),$$

其中 $\tau(l_1^*, l_2^*, \cdots, l_m^*)$ 和 $\tau(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1})$ 分别表示 m 个单项预测方法及再增加一个单项预测方法共 m +1 个单项预测方法参与的基于改进的 Theil 不等系数的组合预测模型的最优目标函数值.

证明 设 $\boldsymbol{L}^* = (l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$ 和 $\bar{\boldsymbol{L}} = (\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_m, \bar{l}_{m+1})^{\mathrm{T}}$ 分别为 m 个单项预测方法参与的组合预测模型 (7.4.5) 的最优解及再增加一个单项预测方法共 m +1 个单项预测方法参与的组合预测模型 (7.4.5) 的最优解, 则有

$$\tau(l_1^*, l_2^*, \cdots, l_m^*) = \sqrt{\boldsymbol{L}^{*\mathrm{T}} \boldsymbol{E} \boldsymbol{L}^*} \Big/ \sqrt{\sum_{t=1}^{N} x_t^2}, \tag{7.4.15}$$

$$\tau(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})=\sqrt{\bar{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}_{m+1}\bar{\boldsymbol{L}}}\Big/\sqrt{\sum_{t=1}^{N}x_t^2},\tag{7.4.16}$$

其中 $\sum_{i=1}^{m}l_i^*=1, l_i^*\geqslant 0, i=1,2,\cdots,m, \sum_{i=1}^{m+1}\bar{l}_i=1, \bar{l}_i\geqslant 0, i=1,2,\cdots,m+1, \boldsymbol{E}_{m+1}=\begin{pmatrix}\boldsymbol{E} & \boldsymbol{a}\\ \boldsymbol{a}^{\mathrm{T}} & E_{(m+1)(m+1)}\end{pmatrix}, \boldsymbol{a}=\left(E_{1(m+1)},E_{2(m+1)},\cdots,E_{m(m+1)}\right)^{\mathrm{T}}$ 为前面的 m 个单项预测方法与第 $(m+1)$ 个单项预测方法预测误差序列的协方差, $E_{(m+1)(m+1)}$ 表示第 $(m+1)$ 个单项预测方法预测误差序列的方差.

令 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*,0)^{\mathrm{T}}$, 显然, $\boldsymbol{L}$ 为 $m+1$ 个单项预测方法参与的组合预测模型 (7.4.5) 的可行解, 则有

$$\tau(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})\leqslant\tau(l_1^*,l_2^*,\cdots,l_m^*,0).$$

由式 (7.4.16) 得

$$\tau(l_1^*,l_2^*,\cdots,l_m^*,0)=\tau(l_1^*,l_2^*,\cdots,l_m^*).$$

所以 $\tau(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})\leqslant\tau(l_1^*,l_2^*,\cdots,l_m^*)$, 结论成立.

定理 7.4.3 证明当再增加一个单项预测方法时, 基于 Theil 不等系数的组合预测模型的最优目标函数值可能不变. 这表明组合预测模型可能存在冗余预测方法.

定理 7.4.4　在基于改进的 Theil 不等系数的组合预测模型中, 若第 i 种单项预测方法优超第 k 种单项预测方法, 则组合预测模型的冗余度至少为 $1/m$.

证明　假设 $\boldsymbol{L}^*=(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型的最优解且第 k 种单项预测方法不为冗余预测方法, 即 $l_k^*>0$, 且 $\sum_{i=1}^{m}l_i^*=1$, 则与 $\boldsymbol{L}^*$ 对应的组合预测模型的目标函数值为

$$\tau(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)=\sqrt{\boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^*}\Big/\sqrt{\sum_{t=1}^{N}x_t^2}.\tag{7.4.17}$$

构造另外一个向量

$$\hat{\boldsymbol{L}}=\boldsymbol{L}^*+\tilde{\boldsymbol{L}}=(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)^{\mathrm{T}},$$

其中 $\tilde{\boldsymbol{L}}=(0,\cdots,l_k^*,\cdots,-l_k^*,\cdots,0)^{\mathrm{T}}$, 即 $\tilde{\boldsymbol{L}}$ 的第 i 个分量为 l_k^*, 第 k 个分量为 $-l_k^*$, 其余 $m-2$ 个分量全为 0. 显然 $\hat{\boldsymbol{L}}$ 为组合预测模型 (7.4.5) 的一可行解, 则与 $\hat{\boldsymbol{L}}$ 对应组合预测模型的目标函数值为

$$\tau(l_1^*,\cdots,l_i^*+l_k^*,\cdots,0,\cdots,l_m^*)=\sqrt{\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\hat{\boldsymbol{L}}}\Big/\sqrt{\sum_{t=1}^{N}x_t^2}.\tag{7.4.18}$$

容易计算

$$\begin{aligned}\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\hat{\boldsymbol{L}} &= (\boldsymbol{L}^* + \tilde{\boldsymbol{L}})^{\mathrm{T}}\boldsymbol{E}(\boldsymbol{L}^* + \tilde{\boldsymbol{L}}) = \boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^* + 2\tilde{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^* + \tilde{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\tilde{\boldsymbol{L}} \\ &= \boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^* + 2l_k^*\sum_{j\neq k} l_j^*(E_{ij} - E_{kj}) + l_k^{*2}(E_{ii} - E_{kk}).\end{aligned} \tag{7.4.19}$$

又因为第 i 种单项预测方法优超第 k 种单项预测方法, 所以由定义 7.4.3 知

$$E_{ii} < E_{kk}, \quad E_{ij} < E_{kj}, \quad j = 1,\cdots,m, \quad j \neq k.$$

注意到假设 $l_k^* > 0$, 且由 (7.4.19) 式知

$$\hat{\boldsymbol{L}}^{\mathrm{T}}\boldsymbol{E}\hat{\boldsymbol{L}} < \boldsymbol{L}^{*\mathrm{T}}\boldsymbol{E}\boldsymbol{L}^*. \tag{7.4.20}$$

所以由式 (7.4.17)、(7.4.18)、(7.4.20) 得

$$\tau(l_1^*,\cdots,l_i^* + l_k^*,\cdots,0,\cdots,l_m^*) < \tau(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*).$$

而这与 $\boldsymbol{L} = (l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型 (7.4.5) 的最优解矛盾! 所以假设不成立. 此即组合预测模型 (7.4.5) 的冗余度至少为 $1/m$.

第 8 章　基于诱导有序信息集结算子的最优组合预测模型及其有效性理论

现有的加权平均组合预测方法存在赋权的缺陷. 本章在有序加权平均算子概念的基础上, 提出诱导有序加权平均算子, 建立新的组合预测模型, 给出组合预测权系数的确定的数学规划方法. 并且进行实例分析, 结果显示该类模型能有效提高组合预测精度.

8.1　三种主要的有序信息集结算子和诱导有序集结算子

8.1.1　OWA 算子和 IOWA 算子的概念及性质[93,105]

美国著名学者 Yager 提出了有序加权平均算子[94,95] (ordered weighted averaging operator) 和诱导有序加权平均算子 [96](induced ordered weighted averaging operator). 有序加权平均算子和诱导有序加权平均算子均是介于最大算子与最小算子之间的一种信息集成方法, 常规的加权算术平均算子是它们的特例. 近年来, 有关该算子的理论研究已引起专家学者的关注, 成为国外十分活跃的研究课题之一, 并广泛应用于数学规划[97]、神经网络[98]、群决策分析[99~102]、模糊逻辑控制器[103]、市场研究[104] 等诸多领域.

定义 8.1.1[93,105]　设 $\mathrm{OWA}_W: R^n \to R$ 为 n 元函数, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 是与 OWA_W 有关的加权向量, 满足 $\sum\limits_{i=1}^{n} w_i=1$, $w_i \geqslant 0$, i=1,2,$\cdots$, n, 若

$$\mathrm{OWA}_W(a_1,a_2,\cdots,a_n)=\sum_{i=1}^{n} w_i b_i. \tag{8.1.1}$$

则称函数OWA_W是 n维有序加权平均算子, 简称为OWA算子, 其中 b_i是 a_1,a_2,$\cdots$,a_n 中按从大到小的顺序排列的第 i 个大的数.

例如, 设 $w_1=0.3$, $w_2=0.4$, $w_3=0.2$, $w_4=0.1$, 则由定义 8.1.1 得

$$\mathrm{OWA}_W(2,4,1,5)=5\times 0.3+4\times 0.4+2\times 0.2+1\times 0.1=3.6.$$

定义 8.1.1 表明, OWA 算子是对 n 个数 $a_1,a_2,\cdots,a_n$ 按从大到小的顺序排序后进行有序加权平均的, 权系数 w_i 与数 a_i 无关, 而是与 $a_1,a_2,\cdots,a_n$ 的按从大小顺序排的第 i 个位置有关.

OWA 算子具有如下性质.

性质 8.1.1(单调性) 设 $(a_1,a_2,\cdots,a_n)$ 和 $(a'_1,a'_2,\cdots,a'_n)$ 是任意两个数据向量, 且有 $a_i \geqslant a'_i$, $\forall i \in \{1,2,\cdots,n\}$, 则

$$\mathrm{OWA}_W(a_1,a_2,\cdots,a_n) \geqslant \mathrm{OWA}_W(a'_1,a'_2,\cdots,a'_n). \tag{8.1.2}$$

证明 由定义 8.1.1 知

$$\mathrm{OWA}_W(a_1,a_2,\cdots,a_n) = \sum_{i=1}^{n} w_i b_i, \tag{8.1.3}$$

其中 b_i 是 $(a_1,a_2,\cdots,a_n)$ 中按从大到小的顺序排列的第 i 个大的数.

$$\mathrm{OWA}_W(a'_1,a'_2,\cdots,a'_n) = \sum_{i=1}^{n} w_i b'_i, \tag{8.1.4}$$

其中 b'_i 是 $a'_1,a'_2,\cdots,a'_n$ 中按从大到小的顺序排列的第 i 个大的数.

因为 $a_i \geqslant a'_i$, $\forall i \in \{1,2,\cdots,n\}$, 则有

$$b_i \geqslant b'_i, \quad \forall i \in \{1,2,\cdots,n\},$$

注意到, $w_i \geqslant 0$, i=1,2,$\cdots$, n, 从而有

$$\sum_{i=1}^{n} w_i b_i \geqslant \sum_{i=1}^{n} w_i b'_i,$$

此即 $\mathrm{OWA}_W(a_1,a_2,\cdots,a_n) \geqslant \mathrm{OWA}_W(a'_1,a'_2,\cdots,a'_n)$. 证毕.

性质 8.1.2(置换不变性) 设 $(a'_1,a'_2,\cdots,a'_n)$ 是 $(a_1,a_2,\cdots,a_n)$ 的任一置换, 则

$$\mathrm{OWA}_W(a'_1,a'_2,\cdots,a'_n) = \mathrm{OWA}_W(a_1,a_2,\cdots,a_n). \tag{8.1.5}$$

证明 设 b_i 和 b'_i 分别是两个数据向量 $a_1,a_2,\cdots,a_n$ 和 $a'_1,a'_2,\cdots,a'_n$ 中按从大到小的顺序排列的第 i 个大的数, 因为 $(a'_1,a'_2,\cdots,a'_n)$ 是 $(a_1,a_2,\cdots,a_n)$ 的任一置换, 所以 $b_i \geqslant b'_i$, $\forall i \in \{1,2,\cdots,n\}$, 从而有

$$\sum_{i=1}^{n} w_i b_i = \sum_{i=1}^{n} w_i b'_i.$$

由式 (8.1.3)、(8.1.4) 知

$$\mathrm{OWA}_W(a'_1,a'_2,\cdots,a'_n) = \mathrm{OWA}_W(a_1,a_2,\cdots,a_n). \text{ 证毕.}$$

性质 8.1.3(幂等性) 设$(a_1, a_2, \cdots, a_n)$是任一数据向量, 若对任意$i \in \{1, 2, \cdots, n\}$, $a_i = a$, 则有

$$\mathrm{OWA}_W(a_1, a_2, \cdots, a_n) = a. \tag{8.1.6}$$

证明 设 b_i 是数据向量 $a_1, a_2, \cdots, a_n$ 按从大到小的顺序排列的第 i 个大的数, 因为 $a_i = a, i = 1, 2, \cdots, n$, 所以 $b_i = a$, 对任意 $i \in \{1, 2, \cdots, n\}$, 注意到 $\sum_{i=1}^{n} w_i = 1$, $w_i \geqslant 0$, i=1,2,$\cdots$, n, 从而有

$$\sum_{i=1}^{n} w_i b_i = a \sum_{i=1}^{n} w_i = a.$$

由式 (8.1.3) 知 $\mathrm{OWA}_W(a_1, a_2, \cdots, a_n) = a$. 证毕.

性质 8.1.4 设 $W^* = (1, 0, \cdots, 0)$, 则

$$\mathrm{OWA}_{W^*}(a_1, a_2, \cdots, a_n) = \max_i\{a_i\}. \tag{8.1.7}$$

证明 设 b_i 是数据向量 $a_1, a_2, \cdots, a_n$ 按从大到小的顺序排列的第 i 个大的数, 则 $b_1 = \max_i\{a_i\}$, 由式 (8.1.1) 知

$$\mathrm{OWA}_{W^*}(a_1, a_2, \cdots, a_n) = 1 \times b_1 + 0 \times b_2 + \cdots + 0 \times b_n = b_1 = \max_i\{a_i\}. \text{ 证毕.}$$

性质 8.1.5 设 $W_* = (0, \cdots, 0, 1)$, 则

$$\mathrm{OWA}_{W_*}(a_1, a_2, \cdots, a_n) = \min_i\{a_i\}. \tag{8.1.8}$$

证明 设 b_i 是数据向量 $a_1, a_2, \cdots, a_n$ 按从大到小的顺序排列的第 i 个大的数, 则

$$b_n = \min_i\{a_i\}.$$

由式 (8.1.1) 知

$$\mathrm{OWA}_{W_*}(a_1, a_2, \cdots, a_n) = 0 \times b_1 + 0 \times b_2 + \cdots + 1 \times b_n = b_n = \min_i\{a_i\}. \text{ 证毕.}$$

性质 8.1.6 设 $W_{\mathrm{AVE}} = \left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$, 则

$$\mathrm{OWA}_{W_{\mathrm{AVE}}}(a_1, a_2, \cdots, a_n) = \frac{1}{n} \sum_{i=1}^{n} a_i. \tag{8.1.9}$$

证明 设 b_i 是数据向量 $a_1, a_2, \cdots, a_n$ 按从大到小的顺序排列的第 i 个大的数, 则

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i.$$

由式 (8.1.1) 知

$$\mathrm{OWA}_{W_{\mathrm{AVE}}}(a_1, a_2, \cdots, a_n) = \frac{1}{n} \times b_1 + \frac{1}{n} \times b_2 + \cdots + \frac{1}{n} \times b_n = \frac{1}{n}\sum_{i=1}^{n} b_i = \frac{1}{n}\sum_{i=1}^{n} a_i.$$证毕.

性质 8.1.7(介值性) 设 $(a_1, a_2, \cdots, a_n)$ 是任一数据向量, 则

$$\mathrm{OWA}_{W^*}(a_1, a_2, \cdots, a_n) \geqslant \mathrm{OWA}_{W}(a_1, a_2, \cdots, a_n) \geqslant \mathrm{OWA}_{W_*}(a_1, a_2, \cdots, a_n). \quad (8.1.10)$$

证明 设 b_i 是数据向量 $a_1, a_2, \cdots, a_n$ 按从大到小的顺序排列的第 i 个大的数, 由式 (8.1.1) 知

$$\mathrm{OWA}_{W}(a_1, a_2, \cdots, a_n) = \sum_{i=1}^{n} w_i b_i.$$

注意到 $\sum_{i=1}^{n} w_i = 1$, $w_i \geqslant 0, i = 1, 2, \cdots, n$, 则有

$$\min_i\{a_i\} = b_n = \sum_{i=1}^{n} w_i b_n \leqslant \sum_{i=1}^{n} w_i b_i \leqslant \sum_{i=1}^{n} w_i b_1 = b_1 = \max_i\{a_i\}.$$

所以由式 (8.1.7)、(8.1.8) 知

$$\mathrm{OWA}_{W_*}(a_1, a_2, \cdots, a_n) \leqslant \mathrm{OWA}_{W}(a_1, a_2, \cdots, a_n) \leqslant \mathrm{OWA}_{W^*}(a_1, a_2, \cdots, a_n).$$ 证毕.

定义 8.1.2[95] 设 $(\langle v_1, a_1\rangle, \langle v_2, a_2\rangle, \cdots, \langle v_n, a_n\rangle)$ 为 n 个二维数组, 令

$$\mathrm{IOWA}_{W}\left(\langle v_1, a_1\rangle, \langle v_2, a_2\rangle, \cdots, \langle v_n, a_n\rangle\right) = \sum_{i=1}^{n} w_i a_{v-\mathrm{index}(i)}, \quad (8.1.11)$$

其中 $\boldsymbol{W} = (w_1, w_2, \cdots, w_n)^{\mathrm{T}}$ 是与 IOWA_W 有关的加权向量, 满足 $\sum_{i=1}^{n} w_i = 1, w_i \geqslant 0$, $i = 1, 2, \cdots, n$, $v-\mathrm{index}(i)$ 是 $v_1, v_2, \cdots, v_n$ 中按从大到小的顺序排列的第 i 个大的数的下标, 则称函数 IOWA_W 是由 $v_1, v_2, \cdots, v_n$ 所产生的 n 维诱导有序加权算术平均算子, 简称为 IOWA 算子, v_i 称为 a_i 的诱导值.

定义 8.1.2 表明, IOWA 算子是对诱导值 $v_1, v_2, \cdots, v_n$ 按从大到小的顺序排序后所对应的 $a_1, a_2, \cdots, a_n$ 中的数进行有序加权平均, w_i 与数 a_i 的大小和位置无关, 而是与其诱导值所在的位置有关.

例如, 设 $\langle 3,4\rangle, \langle 1,2\rangle, \langle 5,1\rangle, \langle 7,0\rangle$ 为 4 个二维数组, 与 IOWA_W 有关的加权向量为

$$w_1 = 0.3, \quad w_2 = 0.4, \quad w_3 = 0.2, \quad w_4 = 0.1.$$

从而

$$\text{IOWA}_W(\langle 3,4\rangle, \langle 1,2\rangle, \langle 5,1\rangle, \langle 7,0\rangle) = 0\times 0.3 + 1\times 0.4 + 4\times 0.2 + 2\times 0.1 = 1.4.$$

IOWA 算子具有如下性质 [105].

性质 8.1.8(单调性) 设 $(\langle v_1,a_1\rangle, \langle v_2,a_2\rangle, \cdots, \langle v_n,a_n\rangle)$ 和 $(\langle v_1,a_1'\rangle, \langle v_2,a_2'\rangle, \cdots, \langle v_n,a_n'\rangle)$ 是任意两个数据向量, 且有 $a_i \geqslant a_i'$, $\forall i\in\{1,2,\cdots,n\}$, 则

$$\text{IOWA}_W(\langle v_1,a_1\rangle, \langle v_2,a_2\rangle, \cdots, \langle v_n,a_n\rangle) \geqslant \text{IOWA}_W(\langle v_1,a_1'\rangle, \langle v_2,a_2'\rangle, \cdots, \langle v_n,a_n'\rangle). \tag{8.1.12}$$

证明 由定义 8.1.2 知

$$\text{IOWA}_W(\langle v_1,a_1\rangle, \langle v_2,a_2\rangle, \cdots, \langle v_n,a_n\rangle) = \sum_{i=1}^{n} w_i a_{v-\text{index}(i)}, \tag{8.1.13}$$

$$\text{IOWA}_W(\langle v_1,a_1'\rangle, \langle v_2,a_2'\rangle, \cdots, \langle v_n,a_n'\rangle) = \sum_{i=1}^{n} w_i a'_{v-\text{index}(i)}. \tag{8.1.14}$$

因为 $a_i \geqslant a_i'$, $\forall i\in\{1,2,\cdots,n\}$, 则有

$$a_{v-\text{index}(i)} \geqslant a'_{v-\text{index}(i)}, \quad \forall i\in\{1,2,\cdots,n\}.$$

注意到 $w_i \geqslant 0$, i=1,2,$\cdots$, n, 从而有 $\sum\limits_{i=1}^{n} w_i a_{v-\text{index}(i)} \geqslant \sum\limits_{i=1}^{n} w_i a'_{v-\text{index}(i)}$.

由式 (8.1.13)、(8.1.14) 知

$$\text{IOWA}_W(\langle v_1,a_1\rangle, \langle v_2,a_2\rangle, \cdots, \langle v_n,a_n\rangle) \geqslant \text{IOWA}_W(\langle v_1,a_1'\rangle, \langle v_2,a_2'\rangle, \cdots, \langle v_n,a_n'\rangle).$$

证毕.

性质8.1.9(置换不变性) 设$(\langle v_1',a_1'\rangle, \langle v_2',a_2'\rangle, \cdots, \langle v_n',a_n'\rangle)$是$(\langle v_1,a_1\rangle, \langle v_2,a_2\rangle, \cdots, \langle v_n,a_n\rangle)$ 任一置换数据向量, 则

$$\text{IOWA}_W(\langle v_1,a_1\rangle, \langle v_2,a_2\rangle, \cdots, \langle v_n,a_n\rangle) = \text{IOWA}_W(\langle v_1',a_1'\rangle, \langle v_2',a_2'\rangle, \cdots, \langle v_n',a_n'\rangle). \tag{8.1.15}$$

证明 因为 $(\langle v_1',a_1'\rangle, \langle v_2',a_2'\rangle, \cdots, \langle v_n',a_n'\rangle)$ 是 $(\langle v_1,a_1\rangle, \langle v_2,a_2\rangle, \cdots, \langle v_n,a_n\rangle)$ 任一置换数据向量, 则有 $a_{v-\text{index}(i)} = a'_{v'-\text{index}(i)}$, $\forall i\in\{1,2,\cdots,n\}$. 从而有

$$\sum_{i=1}^{n} w_i a_{v-\text{index}(i)} = \sum_{i=1}^{n} w_i a'_{v'-\text{index}(i)}.$$

由式 (8.1.13)、(8.1.14) 知

$\mathrm{IOWA}_W(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)=\mathrm{IOWA}_W(\langle v_1,a_1'\rangle,\langle v_2,a_2'\rangle,\cdots,\langle v_n,a_n'\rangle)$.证毕.

性质 8.1.10(幂等性) 设 $(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)$ 是任一数据向量, 若对任意 $i\in\{1,2,\cdots,n\}$, 有 $a_i=a$, 则有

$$\mathrm{IOWA}_W(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)=a. \tag{8.1.16}$$

证明 因为 $a_i=a,\forall i=1,2,\cdots,n$, 所以有 $a_{v-\mathrm{index}(i)}=a$, $\forall i\in\{1,2,\cdots,n\}$, 注意到 $\sum_{i=1}^{n}w_i=1$, $w_i\geqslant 0$, i=1,2,$\cdots$, n, 从而

$\mathrm{IOWA}_W(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)=\sum_{i=1}^{n}w_ia_{v-\mathrm{index}(i)}=\sum_{i=1}^{n}w_ia=a\sum_{i=1}^{n}w_i=a$.证毕.

性质 8.1.11 设 $(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)$ 是任一数据向量, $W_{\mathrm{AVE}}=\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$, 则

$$\mathrm{IOWA}_{W_{\mathrm{AVE}}}(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)=\frac{1}{n}\sum_{i=1}^{n}a_i. \tag{8.1.17}$$

证明 由 $W_{\mathrm{AVE}}=\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$ 和式 (8.1.11) 知

$$\mathrm{IOWA}_{W_{\mathrm{AVE}}}(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)=\sum_{i=1}^{n}\frac{1}{n}a_{v-\mathrm{index}(i)}=\frac{1}{n}\sum_{i=1}^{n}a_{v-\mathrm{index}(i)}.$$

因为$\sum_{i=1}^{n}a_i=\sum_{i=1}^{n}a_{v-\mathrm{index}(i)}$, 所以$\mathrm{IOWA}_{W_{\mathrm{AVE}}}(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)=\frac{1}{n}\sum_{i=1}^{n}a_i$.
证毕.

性质 8.1.12(介值性) 设 $(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)$ 是任一数据向量, 则

$$\min_i\{a_i\}\leqslant\mathrm{IOWA}_W(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)\leqslant\max_i\{a_i\}. \tag{8.1.18}$$

证明 因为$\min_i\{a_i\}\leqslant a_{v-\mathrm{index}(i)}\leqslant\max_i\{a_i\}$, $\forall i\in\{1,2,\cdots,n\}$, 注意到 $\sum_{i=1}^{n}w_i=1$, $w_i\geqslant 0$, i=1,2,$\cdots$, n, 从而

$$\begin{aligned}\mathrm{IOWA}_W(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)&=\sum_{i=1}^{n}w_ia_{v-\mathrm{index}(i)}\\&\leqslant\sum_{i=1}^{n}w_i\max_i\{a_i\}=\max_i\{a_i\}\sum_{i=1}^{n}w_i=\max_i\{a_i\}.\end{aligned}$$

同理可证

$$\min_i\{a_i\} \leqslant \text{IOWA}_W\left(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle\right).\text{证毕.}$$

性质 8.1.13 设 $(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle)$ 是任一数据向量, 若 $v_i=a_i$, $\forall i\in\{1,2,\cdots,n\}$, IOWA 算子退化为 OWA 算子, 即

$$\text{IOWA}_W\left(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle\right)=\text{OWA}_W(a_1,a_2,\cdots,a_n). \tag{8.1.19}$$

证明 因为 $v_i=a_i$, $\forall i\in\{1,2,\cdots,n\}$, $a_{v-\text{index}(i)}=b_i$, 其中 b_i 是 $a_1,a_2,\cdots,a_n$ 中按从大到小的顺序排列的第 i 个大的数, $\forall i\in\{1,2,\cdots,n\}$, 从而

$$\begin{aligned}\text{IOWA}_W\left(\langle v_1,a_1\rangle,\langle v_2,a_2\rangle,\cdots,\langle v_n,a_n\rangle\right)&=\sum_{i=1}^n w_i a_{v-\text{index}(i)}\\&=\sum_{i=1}^n w_i b_i=\text{OWA}_W(a_1,a_2,\cdots,a_n).\end{aligned}$$

证毕.

8.1.2 OWGA 算子和 IOWGA 算子的概念及性质[105]

定义 8.1.3[106] 设 $\text{OWGA}_W: {R^+}^n\to R^+$ 为 n 元函数, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 是与 OWGA_W 有关的加权向量, 满足 $\sum\limits_{i=1}^n w_i=1$, $w_i\geqslant 0$, i=1,2,$\cdots$, n, 若

$$\text{OWGA}_W(a_1,a_2,\cdots,a_n)=\prod_{i=1}^n b_i^{w_i}, \tag{8.1.20}$$

其中 b_i 是 $a_1,a_2,\cdots,a_n$ 中按从大到小的顺序排列的第 i 个大的数. 则称函数 OWGA_W 是 n 维有序加权几何平均算子, 简记为 OWGA 算子.

例如, 设 $w_1=0.3$, $w_2=0.4$, $w_3=0.2$, $w_4=0.1$, 则定义 8.1.3 得

$$\text{OWGA}_W(2,4,1,5)=5^{0.3}\times 4^{0.4}\times 2^{0.2}\times 1^{0.1}=3.2413.$$

定义 8.1.3 表明, OWGA_W 算子是对 n 个数 $a_1,a_2,\cdots,a_n$ 按从大到小的顺序排序后进行有序加权几何平均的, 权系数 w_i 与数 a_i 无关, 而是与 $a_1,a_2,\cdots,a_n$ 的按从大小顺序排的第 i 个位置的数 b_i 有关.

类似地, 可以证明 OWGA 算子具有如下性质[105], 限于篇幅证明省略.

性质 8.1.14(单调性) 设 $(a_1,a_2,\cdots,a_n)$ 和 $(a_1',a_2',\cdots,a_n')$ 是任意两个正的数据向量, 且有 $\forall a_i\geqslant a_i'$, $i\in\{1,2,\cdots,n\}$, 则

$$\text{OWGA}_W(a_1,a_2,\cdots,a_n)\geqslant\text{OWGA}_W(a_1',a_2',\cdots,a_n'). \tag{8.1.21}$$

性质 8.1.15(置换不变性) 设 $(a_1', a_2', \cdots, a_n')$ 是 $(a_1, a_2, \cdots, a_n)$ 的任一置换, 则

$$\mathrm{OWGA}_W(a_1', a_2', \cdots, a_n') = \mathrm{OWGA}_W(a_1, a_2, \cdots, a_n). \tag{8.1.22}$$

性质 8.1.16(幂等性) 设 $(a_1, a_2, \cdots, a_n)$ 是任一正的数据向量, 若对任意 $i \in \{1, 2, \cdots, n\}$, 有 $a_i = a$, 则

$$\mathrm{OWGA}_M(a_1, a_2, \cdots, a_n) = a. \tag{8.1.23}$$

性质 8.1.17 设 $W^* = (1, 0, \cdots, 0), W_* = (0, \cdots, 0, 1)$, 则

$$\mathrm{OWGA}_{W^*}(a_1, a_2, \cdots, a_n) = \max_i\{a_i\}, \tag{8.1.24}$$

$$\mathrm{OWGA}_{W_*}(a_1, a_2, \cdots, a_n) = \min_i\{a_i\}. \tag{8.1.25}$$

性质 8.1.18 设 $W_{\mathrm{AVE}} = \left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$, 则

$$\mathrm{OWGA}_{W_{\mathrm{AVE}}}(a_1, a_2, \cdots, a_n) = \sqrt[n]{\prod_{i=1}^{n} a_i}. \tag{8.1.26}$$

性质 8.1.19(介值性) 设 $(a_1, a_2, \cdots, a_n)$ 是任一正的数据向量, 则

$$\min_i\{a_i\} \leqslant \mathrm{OWGA}_W(a_1, a_2, \cdots, a_n) \leqslant \max_i\{a_i\}. \tag{8.1.27}$$

定义 8.1.4[105] 设 $(\langle u_1, a_1\rangle, \langle u_2, a_2\rangle, \cdots, \langle u_n, a_n\rangle)$ 为 n 个二维数组, 令

$$\mathrm{IOWGA}_W(\langle u_1, a_1\rangle, \langle u_2, a_2\rangle, \cdots, \langle u_n, a_n\rangle) = \prod_{i=1}^{n} a_{u-\mathrm{index}(i)}^{w_i}. \tag{8.1.28}$$

则称函数 IOWGA_W 是由 $u_1, u_2, \cdots, u_n$ 所产生的 n 维诱导有序加权几何平均算子, 简记为 IOWGA 算子, u_i 称为 a_i 的诱导值. 其中 $u-\mathrm{index}(i)$ 是 $u_1, u_2, \cdots, u_n$ 中按从大到小的顺序排列的第 i 个大的数的下标, $\boldsymbol{W} = (w_1, w_2, \cdots, w_n)^{\mathrm{T}}$ 是 OWGA 的加权向量, 满足 $\sum\limits_{i=1}^{n} w_i = 1$, $w_i \geqslant 0$, i=1,2,$\cdots$, n.

定义 8.1.4 表明, IOWGA 算子是对诱导值 $u_1, u_2, \cdots, u_n$ 按从大到小的顺序排序后所对应的 $a_1, a_2, \cdots, a_n$ 中的数进行有序加权几何平均, w_i 与数 a_i 的大小和位置无关, 而是与其诱导值所在的位置有关.

例如, 设 $\langle 3, 4\rangle, \langle 1, 2\rangle, \langle 5, 1\rangle, \langle 7, 5\rangle$ 为 4 个二维数组, OWGA 的加权向量为

$$w_1 = 0.3, \quad w_2 = 0.4, \quad w_3 = 0.2, \quad w_4 = 0.1.$$

则

$$\text{IOWGA}_W(\langle 3,4\rangle,\langle 1,2\rangle,\langle 5,1\rangle,\langle 7,5\rangle) = 5^{0.3}\times 1^{0.4}\times 4^{0.2}\times 2^{0.1} = 2.292.$$

类似地, 可以证明 IOGWA 算子具有如下性质.

性质 8.1.20(单调性) 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 和 $(\langle u_1,a_1'\rangle,\langle u_2,a_2'\rangle,\cdots,\langle u_n,a_n'\rangle)$ 是任意两个正的数据向量, 且有 $a_i \geqslant a_i'$, $\forall i\in\{1,2,\cdots,n\}$, 则

$$\text{IOWGA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle) \geqslant \text{IOWGA}_W(\langle u_1,a_1'\rangle,\langle u_2,a_2'\rangle,\cdots,\langle u_n,a_n'\rangle). \tag{8.1.29}$$

性质 8.1.21(置换不变性) 设 $(\langle u_1',a_1'\rangle,\langle u_2',a_2'\rangle,\cdots,\langle u_n',a_n'\rangle)$ 是 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 任一置换正的数据向量, 则

$$\text{IOWGA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle) = \text{IOWGA}_W(\langle u_1',a_1'\rangle,\langle u_2',a_2'\rangle,\cdots,\langle u_n',a_n'\rangle). \tag{8.1.30}$$

性质 8.1.22(幂等性) 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, 若对任意 $i\in\{1,2,\cdots,n\}$, 有 $a_i=a$, 则有

$$\text{IOWGA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle) = a. \tag{8.1.31}$$

性质 8.1.23 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, $W_{\text{AVE}} = \left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$, 则

$$\text{IOWGA}_{W_{\text{AVE}}}(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle) = \sqrt[n]{\prod_{i=1}^{n} a_i}. \tag{8.1.32}$$

性质 8.1.24(介值性) 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, 则

$$\min_i\{a_i\} \leqslant \text{IOWGA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle) \leqslant \max_i\{a_i\}. \tag{8.1.33}$$

性质 8.1.25 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, 若 $u_i = a_i$, $\forall i\in\{1,2,\cdots,n\}$, IOWGA 算子退化为 OWGA 算子, 即

$$\text{IOWGA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle) = \text{OWGA}_W(a_1,a_2,\cdots,a_n). \tag{8.1.34}$$

8.1.3 OWHA 算子和 IOWHA 算子的概念及性质[107]

定义 8.1.5 设 $\text{OWHA}_W: R^{+^n}\to R^+$ 为 n 元函数, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\text{T}}$ 是与 OWHA_W 有关的加权向量, 满足 $\sum_{i=1}^{n} w_i = 1$, $w_i\geqslant 0$, i=1, 2, $\cdots$,n, 令

$$\text{OWHA}_W(a_1,a_2,\cdots,a_n) = 1\bigg/\sum_{i=1}^{n}(w_i/b_i), \tag{8.1.35}$$

其中 b_i 是 $a_1, a_2, \cdots, a_n$ 中按从大到小的顺序排列的第 i 个大的数, 则称函数 OWHA_W 是 n 维有序加权调和平均算子, 也称为 OWHA 算子.

例如, 设 $w_1 = 0.3$, $w_2 = 0.4$, $w_3 = 0.2$, $w_4 = 0.1$, 则由式 (8.1.35) 得

$$\mathrm{OWHA}_W(2, 4, 1, 5) = 1/(0.3/5 + 0.4/4 + 0.2/2 + 0.1/1) = 2.7778.$$

式 (8.1.35) 表明, OWHA_W 算子是对 n 个数 $a_1, a_2, \cdots, a_n$ 按从大到小的顺序排序后进行有序加权调和平均, 其特点是权系数 w_i 与数 a_i 无关, 而是与 $a_1, a_2, \cdots, a_n$ 按从大小顺序排序的第 i 个位置的数 b_i 有关.

容易证明, OWHA_W 算子具有如下性质.

性质 8.1.26(单调性) 设 $a_1, a_2, \cdots, a_n$ 和 $(a'_1, a'_2, \cdots, a'_n)$ 是两个 n 维向量, 若对任意 i 均有 $a_i \leqslant a'_i$, 则有

$$\mathrm{OWHA}_W(a_1, a_2, \cdots, a_m) \leqslant \mathrm{OWHA}_W(a'_1, a'_2, \cdots, a'_m). \tag{8.1.36}$$

性质 8.1.27(置换不变性) 设 $a_1, a_2, \cdots, a_n$ 是一个 n 维向量, $(a'_1, a'_2, \cdots, a'_n)$ 是 $a_1, a_2, \cdots, a_n$ 任意的一个置换, 则有

$$\mathrm{OWHA}_W(a_1, a_2, \cdots, a_n) = \mathrm{OWHA}_W(a'_1, a'_2, \cdots, a'_n). \tag{8.1.37}$$

性质 8.1.28(幂等性) 设 $a_1, a_2, \cdots, a_n$ 是一个 n 维向量, 若对任意 i, 均有 $a_i = a$, 则有

$$\mathrm{OWHA}_W(a_1, a_2, \cdots, a_n) = a. \tag{8.1.38}$$

性质 8.1.29(介值性) OWHA_W 算子介于 min 算子和 max 算子之间, 即

$$\min_{1\leqslant i\leqslant n}\{a_i\} \leqslant \mathrm{OWHA}_W(a_1, a_2, \cdots, a_n) \leqslant \max_{1\leqslant i\leqslant n}\{a_i\}. \tag{8.1.39}$$

性质 8.1.30 若加权向量 $\boldsymbol{W}_{\mathrm{AVE}} = \left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$, 则此时的 OWHA_W 算子为简单加权调和平均算子, 即

$$\mathrm{OWHA}_W(a_1, a_2, \cdots, a_n) = n \bigg/ \sum_{i=1}^{n} (1/a_i). \tag{8.1.40}$$

文献 [95] 提出诱导有序加权算术平均 (IOWA) 算子的概念, 在此基础上提出如下诱导有序加权调和平均算子.

定义 8.1.6 设 $(\langle u_1, a_1\rangle, \langle u_2, a_2\rangle, \cdots, \langle u_n, a_n\rangle)$ 为 n 个二维数组, $a_i > 0, i = 1, 2, \cdots, n, \boldsymbol{W} = (w_1, w_2, \cdots, w_n)^{\mathrm{T}}$ 是 OWHA_W 的加权向量, 满足 $\sum\limits_{i=1}^{n} w_i = 1, w_i \geqslant$

$0, i=1,2,\cdots,n$, 令

$$\mathrm{IOWHA}_W\left(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle\right)=1\Big/\sum_{i=1}^{n}\frac{w_i}{a_{u-\mathrm{index}(i)}}. \tag{8.1.41}$$

则称函数 IOWHA_W 是由 $u_1,u_2,\cdots,u_n$ 所产生的 n 维诱导有序加权调和平均算子, 也简记为 IOWHA 算子, u_i 称为 a_i 的诱导值. 其中 $u-\mathrm{index}(i)$ 是 $u_1,u_2,\cdots,u_n$ 中按从大到小的顺序排列的第 i 个大的数的下标.

式 (8.1.41) 表明, IOWHA 算子是对诱导值 $u_1,u_2,\cdots,u_n$ 按从大到小的顺序排序后所对应的 $a_1,a_2,\cdots,a_n$ 中的数进行有序加权调和平均权系数 w_i 与数 a_i 的大小和位置无关, 而是与其诱导值所在的位置有关.

例如, 设 $\langle 3,4\rangle,\langle 1,2\rangle,\langle 5,1\rangle,\langle 7,5\rangle$ 为 4 个二维数组, OWHA_W 的加权向量为

$$w_1=0.3,\quad w_2=0.4,\quad w_3=0.2,\quad w_4=0.1,$$

则由式 (8.1.41) 得

$$\mathrm{IOWHA}_W\left(\langle 3,4\rangle,\langle 1,2\rangle,\langle 5,1\rangle,\langle 7,5\rangle\right)=1/(0.3/5+0.4/1+0.2/4+0.1/2)=1.7857.$$

IOWHA 算子具有如下性质 [107].

性质 8.1.31(单调性) 设$(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 和 $(\langle u_1,a_1'\rangle,\langle u_2,a_2'\rangle,\cdots,\langle u_n,a_n'\rangle)$ 是任意两个正的数据向量, 且有 $a_i\geqslant a_i'$, $\forall i\in\{1,2,\cdots,n\}$, 则

$$\mathrm{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)\geqslant\mathrm{IOWHA}_W\left(\langle u_1,a_1'\rangle,\langle u_2,a_2'\rangle,\cdots,\langle u_n,a_n'\rangle\right). \tag{8.1.42}$$

证明 因为$\forall a_i\geqslant a_i'$, $i\in\{1,2,\cdots,n\}$, 则有 $a_{u-\mathrm{index}(i)}\geqslant a'_{u-\mathrm{index}(i)}$, $\forall i\in\{1,2,\cdots,n\}$, 注意到, $w_i\geqslant 0$, $i=1, 2, \cdots, n$, 从而有 $1\Big/\sum_{i=1}^{n}\frac{w_i}{a_{u-\mathrm{index}(i)}}\geqslant 1\Big/\sum_{i=1}^{n}\frac{w_i}{a'_{u-\mathrm{index}(i)}}$. 由定义 8.1.6 知

$\mathrm{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)\geqslant\mathrm{IOWHA}_W(\langle u_1,a_1'\rangle,\langle u_2,a_2'\rangle,\cdots,\langle u_n,a_n'\rangle)$.证毕.

性质 8.1.32(置换不变性) 设 $(\langle u_1',a_1'\rangle,\langle u_2',a_2'\rangle,\cdots,\langle u_n',a_n'\rangle)$ 是 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 任一置换正的数据向量, 则

$$\mathrm{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)=\mathrm{IOWHA}_W(\langle u_1',a_1'\rangle,\langle u_2',a_2'\rangle,\cdots,\langle u_n',a_n'\rangle). \tag{8.1.43}$$

证明 因为 $(\langle u_1',a_1'\rangle,\langle u_2',a_2'\rangle,\cdots,\langle u_n',a_n'\rangle)$ 是 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 任一置换数据向量, 则有 $a_{u-\mathrm{index}(i)}=a'_{u'-\mathrm{index}(i)}$, $\forall i\in\{1,2,\cdots,n\}$, 从而有

$$1\Big/\sum_{i=1}^{n}\frac{w_i}{a_{u-\mathrm{index}(i)}}=1\Big/\sum_{i=1}^{n}\frac{w_i}{a'_{u-\mathrm{index}(i)}}.$$

由定义 8.1.6 知结论成立. 证毕.

性质 8.1.33(幂等性) 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, 若对任意 $i\in\{1,2,\cdots,n\}$, 有 $a_i=a$, 则有

$$\text{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)=a. \tag{8.1.44}$$

证明 因为 $a_i=a, i=1,2,\cdots,n$, 所以有 $a_{u-\text{index}(i)}=a$, $\forall i\in\{1,2,\cdots,n\}$, 注意到 $\sum_{i=1}^{n}w_i=1$, $w_i\geqslant 0$, i=1,2,$\cdots$, n, 从而

$$\text{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)=1\Bigg/\sum_{i=1}^{n}\frac{w_i}{a_{u-\text{index}(i)}}=1\Bigg/\sum_{i=1}^{n}\frac{w_i}{a}=a.$$ 证毕.

性质 8.1.34 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, $\boldsymbol{W}_{\text{AVE}}=\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$, 则

$$\text{IOWHA}_{W_{\text{AVE}}}(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)=n\Bigg/\sum_{i=1}^{n}\frac{1}{a_i}. \tag{8.1.45}$$

证明 由 $\boldsymbol{W}_{\text{AVE}}=\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$ 和式 (8.1.41) 知

$$\text{IOWHA}_{W_{\text{AVE}}}(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)=1\Bigg/\sum_{i=1}^{n}\frac{1/n}{a_{u-\text{index}(i)}}=n\Bigg/\sum_{i=1}^{n}\frac{1}{a_{u-\text{index}(i)}}.$$

因为 $\sum_{i=1}^{n}\frac{1}{a_i}=\sum_{i=1}^{n}\frac{1}{a_{u-\text{index}(i)}}$, 所以有

$$\text{IOWHA}_{W_{\text{AVE}}}(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)=n\Bigg/\sum_{i=1}^{n}\frac{1}{a_i}.$$ 证毕.

性质 8.1.35(介值性) 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, 则

$$\min_i\{a_i\}\leqslant\text{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)\leqslant\max_i\{a_i\}. \tag{8.1.46}$$

证明 因为 $\min_i\{a_i\}\leqslant a_{u-\text{index}(i)}\leqslant\max_i\{a_i\}$, $\forall i\in\{1,2,\cdots,n\}$, 注意到 $\sum_{i=1}^{n}w_i=1$, $w_i\geqslant 0$, i=1,2,$\cdots$, n, 从而

$$\begin{aligned}\text{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)&=1\Bigg/\sum_{i=1}^{n}\frac{w_i}{a_{u-\text{index}(i)}}\leqslant 1\Bigg/\sum_{i=1}^{n}\frac{w_i}{\max_i\{a_i\}}\\&=\max_i\{a_i\}\Bigg/\sum_{i=1}^{n}w_i=\max_i\{a_i\}.\end{aligned}$$

类似可证

$$\min_i\{a_i\} \leqslant \mathrm{IOWHA}_W\left(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle\right).$$

所以结论成立. 证毕.

性质 8.1.36 设 $(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle)$ 是任一正的数据向量, 若 $u_i=a_i, \forall i\in\{1,2,\cdots,n\}$,IOWHA 算子退化为 OWHA 算子, 即

$$\mathrm{IOWHA}_W\left(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle\right)=\mathrm{OWHA}_W(a_1,a_2,\cdots,a_n). \quad (8.1.47)$$

证明 因为 $u_i=a_i, \forall i\in\{1,2,\cdots,n\}$, $a_{u-\mathrm{index}(i)}=b_i$, 其中 b_i 是 $a_1,a_2,\cdots,a_n$ 中按从大到小的顺序排列的第 i 个大的数, $\forall i\in\{1,2,\cdots,n\}$, 从而有

$$\begin{aligned}\mathrm{IOWHA}_W(\langle u_1,a_1\rangle,\langle u_2,a_2\rangle,\cdots,\langle u_n,a_n\rangle) &=1\Big/\sum_{i=1}^{n}\frac{w_i}{a_{u-\mathrm{index}(i)}}\\ &=1\Big/\sum_{i=1}^{n}\frac{w_i}{b_i}=\mathrm{OWHA}_W(a_1,a_2,\cdots,a_n).\end{aligned}$$证毕.

8.2 基于 IOWA 算子的组合预测方法

现有传统的组合预测方法是按着单项预测方法的不同而赋予不同的加权平均系数, 同一个单项预测方法在样本区间上各个时点的加权平均系数是不变的. 然而实际上, 就同一个单项预测方法而言, 它在不同时刻的表现可能不相同, 即在某个时点上预测精度较高, 而在另一时点上预测精度较低. 因此现有的组合预测方法存在与现实不符的缺陷. 本节引进诱导有序加权平均 IOWA 算子, 通过每个单项预测方法在样本区间上各个时点的拟合精度的高低按顺序赋权, 以误差平方和为准则建立了新的组合预测模型, 给出 IOWA 算子组合预测模型权系数的确定方法, 并进行实例对比分析.

8.2.1 基于 IOWA 算子的组合预测模型[108]

设某社会经济现象的指标序列的观察值为 $\{x_t, t=1,2,\cdots,N\}$, 设有 m 种可行的单项预测方法对其进行预测, x_{it} 为第 i 种预测方法第 t 时刻的预测值 (或称拟合值), i=1,2,$\cdots$,m,t=1,2,$\cdots$, N. 设 $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为 m 种单项预测在组合预测中的加权系数, 它满足 $\sum_{i=1}^{m}l_i=1$, $l_i\geqslant 0$, i=1,2,$\cdots$, m.

定义 8.2.1 令

$$\hat{x}_t=\sum_{i=1}^{m}l_ix_{it}, \quad t=1,2,\cdots,N. \quad (8.2.1)$$

则称 $\hat{x}_t$ 为第 t 时刻传统的加权算术平均的组合预测值.

令$a_{it}=\begin{cases}1-|(x_t-x_{it})/x_t|, 当|(x_t-x_{it})/x_t|<1时, \\ 0, \qquad 当|(x_t-x_{it})/x_t|\geqslant 1时,\end{cases}$ $i=1,2,\cdots,m,\quad t=1,2,\cdots,N.$

则 a_{it} 表示第 i 种预测方法第 t 时刻预测精度, 显然 $a_{it}\in[0,1]$. 我们把预测精度 a_{it} 看成预测值 x_{it} 的诱导值, 这样 m 种单项预测方法第 t 时刻预测精度和其对应的在样本区间的预测值就构成了 m 个二维数组 $(\langle a_{1t},x_{1t}\rangle,\langle a_{2t},x_{2t}\rangle,\cdots,\langle a_{mt},x_{mt}\rangle)$, 将 m 种单项预测方法第 t 时刻预测精度序列 $a_{1t},a_{2t},\cdots,a_{mt}$ 按从大到小的顺序排列, 设 $a-\mathrm{index}(it)$ 是第 i 个大的预测精度的下标. 根据定义 8.1.2, 有如下概念.

定义 8.2.2 令

$$\mathrm{IOWA}_{\boldsymbol{L}}\left(\langle a_{1t},x_{1t}\rangle,\langle a_{2t},x_{2t}\rangle,\cdots,\langle a_{mt},x_{mt}\rangle\right)=\sum_{i=1}^{m}l_i x_{a-\mathrm{index}(it)}. \tag{8.2.2}$$

则式 (8.2.2) 称为第 t 时刻由预测精度序列 $a_{1t},a_{2t},\cdots,a_{mt}$ 所产生的 IOWA 算子组合预测值.

显然, 式 (8.2.2) 和式 (8.2.1) 的根本区别在于组合预测的赋权系数与单项预测方法无关, 而是与单项预测方法在各时点上的预测精度的大小密切相关, 这就是基于 IOWA 组合预测的特点.

令 $e_{a-\mathrm{index}(it)}=x_t-x_{a-\mathrm{index}(it)}$, 于是 N 期总的组合预测误差平方和 S 为

$$S=\sum_{t=1}^{N}\left(x_t-\sum_{i=1}^{m}l_i x_{a-\mathrm{index}(it)}\right)^2=\sum_{i=1}^{m}\sum_{j=1}^{m}l_i l_j\left(\sum_{t=1}^{N}e_{a-\mathrm{index}(it)}e_{a-\mathrm{index}(jt)}\right). \tag{8.2.3}$$

因此, 以误差平方和为准则的基于 IOWA 算子的组合预测模型可表示成如下最优化模型[108]

$$\min S(\boldsymbol{L})=\sum_{i=1}^{m}\sum_{j=1}^{m}l_i l_j\left(\sum_{t=1}^{N}e_{a-\mathrm{index}(it)}e_{a-\mathrm{index}(jt)}\right),$$
$$\text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{m}l_i=1,\\ l_i\geqslant 0, i=1,2,\cdots,m.\end{cases} \tag{8.2.4}$$

令 $E_{ij}=\sum_{t=1}^{N}e_{a-\mathrm{index}(it)}e_{a-\mathrm{index}(jt)}, i,j=1,2,\cdots,m$, 则称 $\boldsymbol{E}=(E_{ij})_{m\times m}$ 为 m 阶 IOWA 算子的组合预测误差信息方阵, 因此式 (8.2.4) 可表成矩阵形式

$$\min S(\boldsymbol{L})=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L},$$
$$\text{s.t.}\begin{cases}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}=1,\\ \boldsymbol{L}\geqslant\boldsymbol{0},\end{cases} \tag{8.2.5}$$

其中 $\boldsymbol{R}=(1,1,\cdots,1)^{\mathrm{T}}$, 若不考虑 IOWA 算子的组合预测权重向量 $\boldsymbol{L}$ 的非负性, 则有

$$\begin{aligned} &\min S(\boldsymbol{L})=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L},\\ &\text{s.t.}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}=1. \end{aligned} \tag{8.2.6}$$

8.2.2 基于 IOWA 算子的组合预测模型的求解 [108]

定理 8.2.1 若 IOWA 算子的组合预测误差信息方阵$\boldsymbol{E}$为正定矩阵, 则模型 (8.2.6) 存在唯一最优解

$$\boldsymbol{L}^*=\frac{\boldsymbol{E}^{-1}\boldsymbol{R}}{\boldsymbol{R}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{R}}. \tag{8.2.7}$$

证明 因为 $\boldsymbol{E}$ 为正定矩阵, 从而 $\boldsymbol{E}$ 的逆存在并且 $\boldsymbol{E}^{-1}$ 也为正定矩阵. 对模型 (8.2.6) 构造 Lagrange 函数

$$S(\boldsymbol{L})=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{L}+\lambda(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}-1).$$

其中 λ 为 Lagrange 乘子, 由极值的必要条件, 令

$$\partial S(\boldsymbol{L})/\partial \boldsymbol{L}=\boldsymbol{0},\quad \partial S(\boldsymbol{L})/\partial\lambda=0,$$

即有

$$\begin{cases} 2\boldsymbol{E}\boldsymbol{L}+\lambda\boldsymbol{R}=\boldsymbol{0},\\ \boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}-1=0. \end{cases} \tag{8.2.8}$$

解方程组 (8.2.8) 式得

$$\boldsymbol{L}^*=\frac{\boldsymbol{E}^{-1}\boldsymbol{R}}{\boldsymbol{R}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{R}}.$$

因为 $\partial^2 S(\boldsymbol{L})/\partial \boldsymbol{L}^2=\boldsymbol{E}$ 为正定矩阵, 所以由定理 4.1.2 知, 函数 $S(\boldsymbol{L})$ 为严格凸函数且由极值的充分条件知式 (8.2.7) 为模型 (8.2.6) 的唯一最优解. 证毕.

关于模型 (8.2.5) 和模型 (8.2.6) 的最优解之间的关系有如下结论.

定理 8.2.2 若模型 (8.2.6) 的最优解 $\boldsymbol{L}^*=\frac{\boldsymbol{E}^{-1}\boldsymbol{R}}{\boldsymbol{R}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{R}}$ 是模型 (8.2.5) 可行域的内点或边界点, 则 $\boldsymbol{L}^*$ 也是模型 (8.2.5) 最优解.

证明 设 $\boldsymbol{L}^{**}$ 模型 (8.2.5) 的最优解, 因为 $\boldsymbol{L}^*$ 是模型 (8.2.5) 可行域的内点或边界点, 所以 $\boldsymbol{L}^*$ 是模型 (8.2.5) 的可行解, 则有

$$S(\boldsymbol{L}^*)\geqslant S(\boldsymbol{L}^{**}).$$

同时 $\boldsymbol{L}^{**}$ 模型 (8.2.5) 的最优解, 则 $\boldsymbol{L}^{**}$ 一定满足 $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}^{**}=1$, 从而 $\boldsymbol{L}^{**}$ 也是模型 (8.2.6) 的可行解, 则

$$S(\boldsymbol{L}^*)\leqslant S(\boldsymbol{L}^{**}).$$

从而 $S(\boldsymbol{L}^*) \leqslant S(\boldsymbol{L}^{**})$, 因此 $\boldsymbol{L}^*$ 也是模型 (8.2.5) 最优解. 证毕.

定理 8.2.2 表明若 $\boldsymbol{L}^*$ 满足非负性, 直接用式 (8.2.7) 求模型 (8.2.5) 的最优解. 若不满足非负性, 模型 (8.2.5) 实际上是一个二次规划, 可以利用 Kuhn-Tucker 条件将其转化为线性规划或用现成的最优化软件来求解.

8.2.3 实例分析[85]

为了反映所提出的基于 IOWA 的组合预测模型的有效性, 按照预测效果评价原则, 通常选择预测误差平方和 (SSE)、均方误差 (MSE)、平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE)、均方百分比误差 (MSPE) 等指标作为评价指标体系, 具体计算公式见 4.7. 本节应用文献 [85] 的数据进行实例分析. 指标实际值和各单项预测方法预测值数据如表 8.2.1 所示.

表 8.2.1 指标实际值和各单项预测方法预测值

		1	2	3	4	5	6	7	8	9	10	11	12
	x_t	14.9	18.6	22.2	17.6	19.6	24	31.6	43.7	37	47.2		
例 1	x_{1t}	10	14.9	23.3	26.1	17.5	20.2	26.4	36.8	52.5	38.5		
	x_{2t}	12	15.48	18.95	22.43	25.9	29.38	32.85	36.33	39.8	43.28		
	x_t	57.0	65.4	75.4	82.5	92.8	102.7	119.5	143.8	169.7	201.0	251.2	
例 2	x_{1t}	54.52	62.89	72.54	83.67	96.51	111.32	128.41	148.11	170.84	197.06	227.31	
	x_{2t}	64.68	67.74	68.72	76.61	88.42	104.15	123.79	147.35	174.82	206.21	241.51	
	x_t	11.49	13.06	15.34	20.58	23.28	26.46	27.33	34.22	40.19	53.37	77.79	100.63
例 3	x_{1t}	18.47	14.54	12.84	13.38	16.15	21.16	28.40	37.87	49.58	63.53	79.00	98.12
	x_{2t}	10.03	11.23	15.24	18.67	27.78	26.36	29.67	27.40	42.73	47.36	71.00	109.32

表 8.2.2 计算出两个单项预测方法在各个时点处的预测精度序列, 从而构造出第 t时刻预测精度和其对应的在样本区间的预测值的二维数组$\langle a_{1t}, x_{1t}\rangle, \langle a_{2t}, x_{2t}\rangle, \cdots,$ $\langle a_{mt}, x_{mt}\rangle$.

表 8.2.2 各种单项预测方法预测精度

		1	2	3	4	5	6	7	8	9	10	11	12
例 1	a_{1t}	0.6711	0.8011	0.9505	0.5170	0.8929	0.8417	0.8354	0.8421	0.5811	0.8157		
	a_{2t}	0.8054	0.8323	0.8536	0.7256	0.6786	0.7758	0.9604	0.8314	0.9243	0.9169		
例 2	a_{1t}	0.9565	0.9616	0.9621	0.9858	0.9600	0.9161	0.9254	0.9700	0.9933	0.9804	0.9049	
	a_{2t}	0.8653	0.9899	0.9114	0.9286	0.9528	0.9859	0.9641	0.9753	0.9698	0.9741	0.9614	
例 3	a_{1t}	0.3925	0.8867	0.8370	0.6501	0.6937	0.7997	0.9608	0.8933	0.7664	0.8096	0.9844	0.9751
	a_{2t}	0.8729	0.8599	0.9935	0.9072	0.8067	0.9962	0.9144	0.8007	0.9368	0.8874	0.9127	0.9136

按式 (8.2.2) 计算 IOWA 组合预测值. 下面以例 1 说明计算过程.

$$\text{IOWA}_{\boldsymbol{L}}\left(\langle a_{11}, x_{11}\rangle, \langle a_{21}, x_{21}\rangle\right) = \text{IOWA}_{\boldsymbol{L}}\left(\langle 0.6711, 10\rangle, \langle 0.8054, 12\rangle\right) = 12l_1 + 10l_2,$$

$$\begin{aligned}\text{IOWA}_L\left(\langle a_{12},x_{12}\rangle,\langle a_{22},x_{22}\rangle\right)&=\text{IOWA}_L\left(\langle 0.8011,14.9\rangle,\langle 0.8323,15.48\rangle\right)\\&=15.48l_1+14.9l_2,\end{aligned}$$

$$\text{IOWA}_L\left(\langle a_{13},x_{13}\rangle,\langle a_{23},x_{23}\rangle\right)=23.3l_1+18.95l_2$$

$$\text{IOWA}_L\left(\langle a_{14},x_{14}\rangle,\langle a_{24},x_{24}\rangle\right)=22.43l_1+26.1l_2,$$

$$\text{IOWA}_L\left(\langle a_{15},x_{15}\rangle,\langle a_{25},x_{25}\rangle\right)=17.5l_1+25.9l_2,$$

$$\text{IOWA}_L\left(\langle a_{16},x_{16}\rangle,\langle a_{26},x_{26}\rangle\right)=20.2l_1+29.38l_2,$$

$$\text{IOWA}_L\left(\langle a_{17},x_{17}\rangle,\langle a_{27},x_{27}\rangle\right)=32.85l_1+26.4l_2,$$

$$\text{IOWA}_L\left(\langle a_{18},x_{18}\rangle,\langle a_{28},x_{28}\rangle\right)=36.8l_1+36.33l_2,$$

$$\text{IOWA}_L\left(\langle a_{19},x_{19}\rangle,\langle a_{29},x_{29}\rangle\right)=39.8l_1+52.5l_2,$$

$$\text{IOWA}_L\left(\langle a_{1,10},x_{1,10}\rangle,\langle a_{2,10},x_{2,10}\rangle\right)=43.28l_1+38.5l_2.$$

将其代入到式 (8.2.3) 中, 经整理得如下最优化模型

$$\min S(l_1,l_2)=133.91l_1^2+302.83l_1l_2+586.4438l_2^2,$$

$$\text{s.t.}\begin{cases}l_1+l_2=1,\\l_1\geqslant 0,l_2\geqslant 0.\end{cases}$$

利用 Matlab 最优化工具箱计算得基于IOWA算子的组合预测模型的最优权系数为

$$l_1^*=1,\quad l_2^*=0.$$

例 1 的组合预测模型的 IOWA 最优权系数表明, 未来的组合预测是把两个单项预测方法中的预测精度的最高者预测值作为它们的组合预测值. 这可以从表 8.2.1 中得到解释, 就例 1 而言, 有 60%的年份实际值不在两个单项预测方法的预测值所构成的区间里, 即两个单项预测方法的预测值同时高于或同时低于实际值, 因此这两个单项预测方法的信息互补性较差. 所以就取预测精度最高的预测值为其的组合预测值, 否则, 若 IOWA 的组合预测模型的最优权系数都大于零, 这两个单项预测方法预测值的线性组合会偏离实际值更远. 因此这个结论是有道理的.

例 2 的组合预测模型的 IOWA 最优权系数与例 1 相同. 例 3 的组合预测模型的 IOWA 最优权系数是 $l_1^*=0.7095$, $l_2^*=0.2905$. 这个结果与例 1 不同, 原因是例 3 中有 60%的年份的实际值介于两个单项预测方法的预测值之间, 因此两个单项预测方法的信息互补性较强.

另外要说明一点是, 若在传统的组合预测模型中出现某个单项预测方法在组合预测中的最优权系数为 1, 其他单项预测方法在组合预测中的最优权系数为 0, 其意

思是指组合预测值就是某个单项预测方法的预测值, 其他单项预测方法为冗余预测方法, 不提供任何有效信息. 显然, 即使传统的模型和本节提出的基于 IOWA 的组合预测模型 (8.2.4) 获得相同的最优权系数, 它们的实际含义不同.

例 2 和例 3 的计算过程和例 1 是相同的, 所有的 IOWA 的组合预测模型的最优权系数以及预测效果评价的五个误差指标的计算结果如表 8.2.3 所示.

表 8.2.3 预测效果评价指标体系

预测效果评价指标体系			SSE	MAE	MSE	MAPE	MSPE
例 1	单项预测	方法 1	520.60	6.04	2.28	0.2251	0.0825
		方法 2	199.76	4.11	1.41	0.1696	0.0599
	组合预测最优权重向量	$l_1^*=1, l_2^*=0$	137.28	3.37	1.17	0.1308	0.0468
例 2	单项预测	方法 1	795.59	5.78	2.56	0.0440	0.0156
		方法 2	338.25	4.96	1.67	0.0474	0.0179
	组合预测最优权重向量	$l_1^*=1, l_2^*=0$	173.74	3.18	1.20	0.0259	0.0087
例 3	单项预测	方法 1	401.56	4.88	1.67	0.1959	0.0731
		方法 2	245.58	3.59	1.30	0.0998	0.0334
	组合预测最优权重向量	l_1^*=0.7095,l_2^*=0.2905	62.80	1.96	0.66	0.0848	0.0287

从表 8.2.3 预测效果评价指标体系来看, 本节提出的基于 IOWA 的组合预测模型的各种误差指标值均明显的低于文献 [85] 的计算结果, 从而表明所提出的组合预测方法优于传统的组合预测方法, 能够有效地提高预测精度.

基于诱导有序加权平均算子的概念, 本节提出新的组合预测模型, 通过一个二次规划模型求解可以获得样本区间上组合预测 IOWA 最优权系数, 设为 $\boldsymbol{L}^*=(l_1^*, l_2^*, \cdots, l_m^*)^{\mathrm{T}}$, 根据预测连贯性的原则, 可以用它来进行预测区间 $[N+1, N+2, \cdots]$ 的 IOWA 组合预测, 公式为

$$\mathrm{IOWA}_{\boldsymbol{L}^*}(\langle a_{1t}, x_{1t}\rangle, \langle a_{2t}, x_{2t}\rangle, \cdots, \langle a_{mt}, x_{mt}\rangle)=\sum_{i=1}^{m} l_i^* x_{a\text{-index}(it)}, \quad t=N+1, N+2, \cdots,$$

其中在预测区间 $[N+1, N+2, \cdots]$ 上预测精度序列 $a_{1t}, a_{2t}, \cdots, a_{mt}$ 的大小确定原则是依据各个单项预测方法在样本区间上近几期拟合平均精度的高低. 即若要进行未来 k 步的预测, 用第 i 种预测方法最近 k 期拟合平均精度 $\dfrac{1}{k}\sum\limits_{t=N-k+1}^{N} a_{it}$ 来反映预测区间上 $N+k$ 期的预测精度.

8.3 基于 IOWGA 算子的组合预测方法

本节针对现有的加权几何平均组合预测方法, 在诱导有序加权平均 (IOWA) 算子概念的基础上, 提出诱导有序几何加权平均 (IOWGA) 算子, 建立新的组合预测

模型. 该模型赋权的基本思想是依据每个单项预测方法在各个时点的拟合精度的高低按顺序赋权.

8.3.1 基于 IOWGA 算子的组合预测模型的建立[109]

设 $\{x_t, t=1,2,\cdots,N\}$ 为指标序列的实际值, x_{it} 为第 i 种预测方法第 t 时刻的预测值, i=1,2,$\cdots$, m, t=1,2,$\cdots$, N. 设 $l_1,l_2,\cdots,l_m$ 为 m 种单项预测在组合预测中的加权系数, 它满足归一性和非负性, 在定义 8.1.4 的基础上, 考虑如下的 IOWGA 组合预测值模型. 令

$$p_{it}=\begin{cases}1-|(x_t-x_{it})/x_t|, & \text{当}|(x_t-x_{it})/x_t|<1\text{时},\\ 0, & \text{当}|(x_t-x_{it})/x_t|\geqslant 1\text{时}.\end{cases}$$

则 p_{it} 表示第 i 种预测方法第 t 时刻预测精度, 显然 $p_{it}\in[0,1]$. 我们把预测精度 p_{it} 看成预测值 x_{it} 的诱导值, 这样 m 种单项预测方法第 t 时刻预测精度和其对应的在样本区间的预测值就构成了 m 个二维数组 $\langle p_{1t},x_{1t}\rangle,\langle p_{2t},x_{2t}\rangle,\cdots,\langle p_{mt},x_{mt}\rangle$, 设 $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为各种预测方法在组合预测中的 OWGA 的加权向量, 将 m 种单项预测方法第 t 时刻预测精度序列 $p_{1t},p_{2t},\cdots,p_{mt}$ 按从大到小的顺序排列, 设 $p-\mathrm{index}(it)$ 是第 i 个大的预测精度的下标, 根据定义 8.1.4, 则有

定义 8.3.1 令

$$\mathrm{IOWGA}_{\boldsymbol{L}}\left(\langle p_{1t},x_{1t}\rangle,\langle p_{2t},x_{2t}\rangle,\cdots,\langle p_{mt},x_{mt}\rangle\right)=\prod_{i=1}^{m}x_{p-\mathrm{index}(it)}^{l_i},\quad t=1,2,\cdots,N. \tag{8.3.1}$$

则称式 (8.3.1) 为由预测精度序列 $p_{1t},p_{2t},\cdots,p_{mt}$ 所产生的第 t 时刻 IOWGA 组合预测值.

式 (8.3.1) 给出了基于 IOWGA 组合预测的赋权特点. 新的组合预测的赋权系数与单项预测方法种类无关, 而是与各个单项预测方法在各时点上的预测精度的大小密切相关, 预测精度成为第 i 种单项预测方法第 t 时刻的赋权系数的诱导因素. 预测精度高要优先赋大一点的权系数, 这种赋权思想比较符合实际.

为方便求解, 可取对数误差平方和作为优化准则.

令 $e_{a-\mathrm{index}(it)}=\ln x_t-\ln x_{p-\mathrm{index}(it)}$, 于是 N 期总的组合预测对数误差平方和 S 为

$$\begin{aligned}G&=\sum_{t=1}^{N}\left(\ln x_t-\ln\prod_{i=1}^{m}x_{p-\mathrm{index}(it)}^{l_i}\right)^2=\sum_{t=1}^{N}\left(\ln x_t-\sum_{i=1}^{m}l_i\ln x_{a-\mathrm{index}(it)}\right)^2\\&=\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_j\left(\sum_{t=1}^{N}e_{a-\mathrm{index}(it)}e_{a-\mathrm{index}(jt)}\right).\end{aligned} \tag{8.3.2}$$

因此, 以对数误差平方和为准则的基于 IOWGA 算子的最优组合预测模型可表示为[109]

$$\min G(\boldsymbol{L}) = \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j \left(\sum_{t=1}^{N} e_{a-\text{index}(it)} e_{a-\text{index}(jt)}\right),$$

$$\text{s.t}\begin{cases} \sum_{i=1}^{m} l_i = 1, \\ l_i \geqslant 0, i = 1, 2, \cdots, m. \end{cases} \tag{8.3.3}$$

模型 (8.3.3) 实际上是一个二次规划, 可以利用 Kuhn-Tucker 条件将其转化为线性规划或用 Lingo 软件来求解.

8.3.2 实例分析[82]

本节利用文献 [82] 的数据进行实例分析. 某省有 11 个年份社会商品零售总额的统计资料, 现采用两种单项预测方法进行预测, 分别是指数曲线预测方法和抛物线预测方法. 原始统计资料和这两种单项预测模型的预测数值见表 8.3.1.

表 8.3.1 社会商品零售总额实际值及两种单项预测方法预测值

年份 t	零售总额 (亿元)	指数曲线预测值x_{1t}	抛物线预测值x_{2t}	指数曲线预测精度p_{1t}	抛物线预测精度p_{2t}	IOWGA 组合预测值$\hat{x}_t$
1	57.0	54.52	64.48	0.9565	0.8688	56.00
2	65.4	62.89	66.74	0.9616	0.9795	66.11
3	75.4	72.54	68.72	0.9621	0.9114	71.92
4	82.5	83.67	76.61	0.9858	0.9286	82.50
5	92.8	96.51	88.42	0.9600	0.9528	95.17
6	102.7	111.32	104.15	0.9161	0.9859	105.26
7	119.5	128.41	123.79	0.9254	0.9641	124.51
8	143.8	148.11	147.35	0.9700	0.9753	147.47
9	169.7	170.84	174.82	0.9933	0.9698	171.47
10	201.0	197.06	206.21	0.9804	0.9741	198.49
11	215.2	227.31	241.51	0.9437	0.8777	229.52

表 8.3.1 中还计算出两个单项预测方法在各个时点处的预测精度序列, 从而构造出第 t 时刻预测精度和其对应的在样本区间的预测值的二维数组 $\langle p_{1t}, x_{1t}\rangle, \langle p_{2t}, x_{2t}\rangle$, $t = 1, 2, \cdots, 11$. 按式 (8.3.1) 计算 IOWGA 组合预测值.

$$\begin{aligned}\text{IOWGA}_{\boldsymbol{L}}\left(\langle p_{11}, x_{11}\rangle, \langle p_{21}, x_{21}\rangle\right) &= \text{IOWGA}_{\boldsymbol{L}}\left(\langle 0.9565, 54.52\rangle, \langle 0.8688, 64.48\rangle\right) \\ &= 54.52^{l_1} 64.48^{l_2},\end{aligned}$$

$$\text{IOWGA}_{\boldsymbol{L}}\left(\langle p_{12}, x_{12}\rangle, \langle p_{22}, x_{22}\rangle\right) = 66.74^{l_1} 62.89^{l_2},$$

$$\text{IOWGA}_{\boldsymbol{L}}\left(\langle p_{13}, x_{13}\rangle, \langle p_{23}, x_{23}\rangle\right) = 72.54^{l_1} 68.72^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{14},x_{14}\rangle,\langle p_{24},x_{24}\rangle\right)=83.67^{l_1}76.61^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{15},x_{15}\rangle,\langle p_{25},x_{25}\rangle\right)=96.51^{l_1}88.42^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{16},x_{16}\rangle,\langle p_{26},x_{26}\rangle\right)=104.15^{l_1}111.32^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{17},x_{17}\rangle,\langle p_{27},x_{27}\rangle\right)=123.79^{l_1}128.41^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{18},x_{18}\rangle,\langle p_{28},x_{28}\rangle\right)=147.35^{l_1}148.11^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{19},x_{19}\rangle,\langle p_{29},x_{29}\rangle\right)=170.84^{l_1}174.82^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{1,10},x_{1,10}\rangle,\langle p_{2,10},x_{2,10}\rangle\right)=197.06^{l_1}206.21^{l_2},$$

$$\mathrm{IOWGA}_{L}\left(\langle p_{1,11},x_{1,11}\rangle,\langle p_{2,11},x_{2,11}\rangle\right)=227.31^{l_1}241.51^{l_2},$$

其中 l_1,l_2 为两种预测方法在组合预测中的 OWGA 的加权向量. 将其代入到式 (8.3.3) 中, 经计算得如下最优化模型.

$$\min S(l_1,l_2)=0.0111l_1^2+0.0096l_1l_2+0.06058l_2^2,$$

$$\text{s.t.}\begin{cases} l_1+l_2=1,\\ l_1\geqslant 0, l_2\geqslant 0.\end{cases}$$

利用 Matlab 最优化工具箱计算得基于 IOWGA 的组合预测模型的最优权系数为

$$l_1^*=0.8406,\quad l_2^*=0.1594.$$

把最优权系数再代入到 $\mathrm{IOWGA}_{L}\left(\langle p_{1t},x_{1t}\rangle,\langle p_{2t},x_{2t}\rangle\right)$, $t=1,2,\cdots,11$ 中可计算出 IOWGA 的组合预测值, 见表 8.3.1.

为了反映所提出的基于 IOWGA 的组合预测模型的有效性, 按照预测效果评价原则, 通常选择下列指标作为评价指标体系, 经计算

(1) 平方和误差 SSE$=\sum_{t=1}^{11}(x_t-\hat{x}_t)^2=278.83$;

(2) 均方误差 MSE$=\frac{1}{11}\sqrt{\sum_{t=1}^{11}(x_t-\hat{x}_t)^2}=1.5180$;

(3) 平均绝对误差 MAE$=\frac{1}{11}\sum_{t=1}^{11}|x_t-\hat{x}_t|=3.40$;

(4) 平均绝对百分比误差 MAPE$=\frac{1}{11}\sum_{t=1}^{11}|(x_t-\hat{x}_t)/x_t|=0.0256$;

(5) 均方百分比误差 MSPE$=\frac{1}{11}\sqrt{\sum_{t=1}^{11}[(x_t-\hat{x}_t)/x_t]^2}=0.0095$.

文献 [82] 只给出两个预测效果的评价指标, 即平方和误差 SSE=449.41, 均方百分比误差 MSPE=0.0122. 可见就这两个评价指标而言, 本节提出的模型 (8.3.1) 的结果显著地优于文献 [82] 计算结果, 从而表明本节提出的组合预测方法的有效性.

根据预测的原则, 可以用它来进行预测区间 $[N+1, N+2, \cdots]$ 的 IOWGA 组合预测, 若要进行未来 k 步的预测, 用第 i 种预测方法最近 k 期拟合平均精度 $\frac{1}{k}\sum\limits_{t=N-k+1}^{N} p_{it}$ 来反映预测区间上 $N+k$ 期的预测精度的大小. 公式为

$$\mathrm{IOWGA}_{L^*}(\langle p_{1t}, x_{1t}\rangle, \langle p_{2t}, x_{2t}\rangle, \cdots, \langle p_{mt}, x_{mt}\rangle) = \prod_{i=1}^{m} x_{p-\mathrm{index}(it)}^{l_i^*}, \quad t = N+1, N+2, \cdots,$$

其中在预测区间 $[N+1, N+2, \cdots]$ 上预测精度序列 $p_{1t}, p_{2t}, \cdots, p_{mt}$ 的大小确定原则是依据各个单项预测方法在样本区间上近期拟合平均精度的高低.

8.4 IOWHA 算子及其在组合预测中的应用

目前, 理论研究和实际应用较多的是有序加权算术平均 (OWA) 算子[94,95,110] 和有序加权几何平均 (OWGA) 算子[105,106]. 本节在此基础上, 提出有序加权调和平均 (OWHA) 算子, 探讨 OWHA 的性质, 并且提出诱导有序加权调和平均 (IOWHA) 算子的概念, 建立新的基于 IOWHA 算子的组合预测模型.

8.4.1 基于 IOWHA 算子的组合预测模型[107]

设 p_{it} 表示第 i 种预测方法第 t 时刻预测精度, x_{it} 为第 i 种预测方法第 t 时刻的预测值, i=1,2,$\cdots$, m, t=1,2,$\cdots$, N. $\{x_t, t = 1, 2, \cdots, N\}$ 为经济现象的指标序列的观察值, 设 $\boldsymbol{L} = (l_1, l_2, \cdots, l_m)^{\mathrm{T}}$ 为 m 种单项预测在组合预测中的加权系数, 它满足归一性和非负性. 为了克服现有的加权调和平均组合预测方法赋权的缺陷, 本节在定义 8.1.6 的基础上, 建立基于 IOWHA 算子的组合预测模型. 该模型赋权的基本思想是依据每个单项预测方法在各个时点的预测精度的高低按顺序赋权.

定义 8.4.1 令

$$\mathrm{IOWHA}_{L}(\langle p_{1t}, x_{1t}\rangle, \langle p_{2t}, x_{2t}\rangle, \cdots, \langle p_{mt}, x_{mt}\rangle) = 1\Bigg/\sum_{i=1}^{m}\frac{l_i}{x_{p-\mathrm{index}(it)}}, \quad t = 1, 2, \cdots, N. \tag{8.4.1}$$

则式 (8.4.1) 称为由预测精度序列 $p_{1t}, p_{2t}, \cdots, p_{mt}$ 所产生的第 t 时刻 IOWHA 组合预测值. 其中 $p-\mathrm{index}(it)$ 是第 i 个大的预测精度的下标.

令 $e_{a-\mathrm{index}(it)} = 1/x_t - 1/x_{p-\mathrm{index}(it)}$, 于是, N 期总的基于 IOWHA 的组合预

测倒数误差平方和 F 为

$$F=\sum_{t=1}^{N}(1/x_t-1/\hat{x}_t)^2=\sum_{t=1}^{N}\left(\sum_{i=1}^{m}l_i(1/x_t-1/x_{p-\text{index}(it)})\right)^2$$
$$=\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_j\left(\sum_{t=1}^{N}e_{a-\text{index}(it)}e_{a-\text{index}(jt)}\right). \tag{8.4.2}$$

因此, 基于 IOWHA 的组合预测模型可表示成如下模型[107]

$$\min F(\boldsymbol{L})=\sum_{i=1}^{m}\sum_{j=1}^{m}l_il_j\left(\sum_{t=1}^{N}e_{a-\text{index}(it)}e_{a-\text{index}(jt)}\right),$$

$$\text{s.t.}\begin{cases}\sum_{i=1}^{m}l_i=1,\\ l_i\geqslant 0, i=1,2,\cdots,m.\end{cases} \tag{8.4.3}$$

8.4.2 基于 IOWHA 算子的组合预测模型实例分析[111]

为了反映本节提出的基于 IOWHA 的组合预测模型的有效性, 按照预测效果评价原则, 通常选择下列指标作为评价指标体系: (1) 平方和误差; (2) 均方误差; (3) 平均绝对误差; (4) 平均绝对百分比误差; (5) 均方百分比误差.

本节应用文献 [111] 的数据进行实例分析. 指标实际值和各单项预测方法预测值数据如表 8.4.1 所示.

表 8.4.1 指标实际值和各单项预测方法预测值

		1	2	3	4	5	6	7	8	9	10	11	12
	x_t	14.9	18.6	22.2	17.6	19.6	24	31.6	43.7	37	47.2		
例 1	x_{1t}	10	14.9	23.3	26.1	17.5	20.2	26.4	36.8	52.5	38.5		
	x_{2t}	12	15.48	18.95	22.43	25.9	29.38	32.85	36.33	39.8	43.28		
	x_t	57.0	65.4	75.4	82.5	92.8	102.7	119.5	143.8	169.7	201.0	251.2	
例 2	x_{1t}	54.52	62.89	72.54	83.67	96.51	111.32	128.41	148.11	170.84	197.06	227.31	
	x_{2t}	64.68	64.74	68.72	76.61	88.42	104.15	123.79	147.35	174.82	206.21	241.51	
	x_t	11.49	13.06	15.34	20.58	23.28	26.46	27.33	34.22	40.19	53.37	77.79	100.63
例 3	x_{1t}	18.47	14.54	12.84	13.38	16.15	21.16	28.40	37.87	49.58	63.53	79.00	98.12
	x_{2t}	10.03	11.23	15.24	18.67	27.78	26.36	29.67	27.40	42.73	47.36	71.00	109.32

表 8.4.2 计算出两个单项预测方法在各个时点处的预测精度序列. 显然, 同一种单项预测方法在不同时点上预测精度“时高时低”.

表 8.4.2 各种单项预测方法预测精度

		1	2	3	4	5	6	7	8	9	10	11	12
例 1	a_{1t}	0.6711	0.8011	0.9505	0.5170	0.8929	0.8417	0.8354	0.8421	0.5811	0.8157		
	a_{2t}	0.8054	0.8323	0.8536	0.7256	0.6786	0.7758	0.9604	0.8314	0.9243	0.9169		
例 2	a_{1t}	0.9565	0.9616	0.9621	0.9858	0.9600	0.9161	0.9254	0.9700	0.9933	0.9804	0.9049	
	a_{2t}	0.8653	0.9899	0.9114	0.9286	0.9528	0.9859	0.9641	0.9753	0.9698	0.9741	0.9614	
例 3	a_{1t}	0.3925	0.8867	0.8370	0.6501	0.6937	0.7997	0.9608	0.8933	0.7664	0.8096	0.9844	0.9751
	a_{2t}	0.8729	0.8599	0.9935	0.9072	0.8067	0.9962	0.9144	0.8007	0.9368	0.8874	0.9127	0.9136

下面以例 2 说明计算过程. 由表 8.4.1 和表 8.4.2 得到第 t 时刻预测精度和其对应的在样本区间的预测值所构成的二维数组 $\langle p_{1t}, x_{1t}\rangle, \langle p_{2t}, x_{2t}\rangle$, $t=1,2,\cdots,11$. 按式 (8.4.1) 计算 IOWHA 组合预测值如下

$$\begin{aligned}\mathrm{IOWHA}_L\left(\langle p_{11}, x_{11}\rangle, \langle p_{21}, x_{21}\rangle\right) &= \mathrm{IOWHA}\left(\langle 0.9565, 54.52\rangle, \langle 0.8688, 64.48\rangle\right)\\ &= 1/(l_1/54.52 + l_2/64.48).\end{aligned}$$

同理可得

$$\mathrm{IOWHA}_L\left(\langle p_{12}, x_{12}\rangle, \langle p_{22}, x_{22}\rangle\right) = 1/(l_1/66.74 + l_2/62.89),$$

$$\mathrm{IOWHA}_L\left(\langle p_{13}, x_{13}\rangle, \langle p_{23}, x_{23}\rangle\right) = 1/(l_1/72.54 + l_2/68.72),$$

$$\mathrm{IOWHA}_L\left(\langle p_{14}, x_{14}\rangle, \langle p_{24}, x_{24}\rangle\right) = 1/(l_1/83.67 + l_2/76.61),$$

$$\mathrm{IOWHA}_L\left(\langle p_{15}, x_{15}\rangle, \langle p_{25}, x_{25}\rangle\right) = 1/(l_1/96.51 + l_2/88.42),$$

$$\mathrm{IOWHA}_L\left(\langle p_{16}, x_{16}\rangle, \langle p_{26}, x_{26}\rangle\right) = 1/(l_1/104.15 + l_2/111.32),$$

$$\mathrm{IOWHA}_L\left(\langle p_{17}, x_{17}\rangle, \langle p_{27}, x_{27}\rangle\right) = 1/(l_1/123.79 + l_2/128.41),$$

$$\mathrm{IOWHA}_L\left(\langle p_{18}, x_{18}\rangle, \langle p_{28}, x_{28}\rangle\right) = 1/(l_1/147.35 + l_2/148.11),$$

$$\mathrm{IOWHA}_L\left(\langle p_{19}, x_{19}\rangle, \langle p_{29}, x_{29}\rangle\right) = 1/(l_1/170.84 + l_2/174.82),$$

$$\mathrm{IOWHA}_L\left(\langle p_{1,10}, x_{1,10}\rangle, \langle p_{2,10}, x_{2,10}\rangle\right) = 1/(l_1/197.06 + l_2/206.21),$$

$$\mathrm{IOWHA}_L\left(\langle p_{1,11}, x_{1,11}\rangle, \langle p_{2,11}, x_{2,11}\rangle\right) = 1/(l_1/227.31 + l_2/241.51),$$

其中 l_1, l_2 为两种预测方法在组合预测中的 OWHA 的加权向量, 取倒数误差, 将其代入到式 (8.4.2) 中, 经计算得如下最优化模型

$$\min F(l_1, l_2) = (13.024 l_1^2 - 18.1324 l_1 l_2 + 86.945 l_2^2) \times 10^{-7},$$

$$\text{s.t.} \begin{cases} l_1 + l_2 = 1, \\ l_1 \geqslant 0, l_2 \geqslant 0. \end{cases}$$

利用 Matlab 最优化工具箱计算得基于 IOWGA 的组合预测模型的最优权系数为

$$l_1^* = 0.8130, \quad l_2^* = 0.1870.$$

例 1 和例 3 的计算过程和例 2 是相同的, 所有的 IOWHA 的组合预测模型的最优权系数以及预测效果评价的五个误差指标的计算结果如表 8.4.3 所示.

表 8.4.3 预测效果评价指标体系

预测效果评价指标体系			SSE	MAE	MSE	MAPE	MSPE
例 1	单项	方法 1	520.60	6.04	2.28	0.2251	0.0825
	预测	方法 2	199.76	4.11	1.41	0.1696	0.0599
	IOWHA 组合预测	$l_1^*=1$					
	最优权重向量	$l_2^*=0$	137.28	3.37	1.17	0.1308	0.0468
例 2	单项	方法 1	795.59	5.78	2.56	0.0440	0.0156
	预测	方法 2	338.25	4.96	1.67	0.0474	0.0179
	IOWHA 组合预测	$l_1^*=0.8130$					
	最优权重向量	$l_2^*=0.1870$	230.948	3.2696	1.3815	0.0246	0.0087
例 3	单项	方法 1	401.56	4.88	1.67	0.1959	0.0731
	预测	方法 2	245.58	3.59	1.30	0.0998	0.0334
	IOWHA 组合预测	$l_1^*=0.7899$					
	最优权重向量	$l_2^*=0.2101$	43.50	1.4642	0.5496	0.0517	0.0190

从表 8.4.3 预测效果评价指标体系来看, 基于 IOWHA 的组合预测模型不仅大大低于各单项预测方法的误差指标, 而且也明显地低于文献 [111] 的计算结果, 从而表明所提出的组合预测方法优于传统的组合预测方法, 能够有效地提高预测精度. 根据预测的原则, 可以用它来进行预测区间 $[N+1, N+2, \cdots]$ 的 IOWHA 组合预测, 公式为

$$\mathrm{IOWHA}(\langle p_{1t}, x_{1t}\rangle, \langle p_{2t}, x_{2t}\rangle, \cdots, \langle p_{mt}, x_{mt}\rangle) = 1 \Big/ \sum_{i=1}^{m} \frac{l_i^*}{x_{p\text{-index}(it)}}, \quad t = N+1, N+2, \cdots,$$

其中在预测区间 $[N+1, N+2, \cdots]$ 上预测精度序列 $p_{1t}, p_{2t}, \cdots, p_{mt}$ 的大小确定原则是依据各个单项预测方法在样本区间上近期拟合平均精度的高低. 即若要进行未来 k 步的预测, 用第 i 种预测方法最近 k 期拟合平均精度 $\frac{1}{k}\sum_{t=N-k+1}^{N} p_{it}$ 来反映预测区间上 $N+k$ 期的预测精度的大小.

8.5 一类基于 OWA 算子的组合预测模型及其性质

本节引进有序加权平均 (OWA) 算子建立相应的组合预测模型, 指出最大绝对

误差最小化的组合预测模型是其特例, 并给出其线性规划的求解方法. 然后对该模型重新提出若干新概念, 探讨了非劣性组合预测和冗余预测方法的判定.

8.5.1 基于 OWA 算子的组合预测模型的建立[151]

设对同一预测对象的某个指标序列 $\{x_t, t=1,2,\cdots,n\}$, 存在 m 种可行单项预测方法对其进行预测, 设第 i 种单项预测方法在第 t 时刻的预测值为 x_{it}, $i=1,2,\cdots,m$, $t=1,2,\cdots,n$, 设 $l_1,l_2,\cdots,l_m$ 分别为 m 种单项预测方法的加权系数, 为了使组合预测保持无偏性, 加权系数应满足 $\sum\limits_{i=1}^{m} l_i=1$, $l_i\geqslant 0$, $i=1,2,\cdots,m$. 根据加权算术平均的组合预测原理, 则有

$$\hat{x}_t=\sum_{i=1}^{m} l_i x_{it},\quad t=1,2,\cdots,n,$$

其中 $\hat{x}_t$ 表示第 t 时刻预测对象实际值 x_t 的组合预测值.

设 e_t 为组合预测在第 t 时刻的预测误差, 称 $e_{it}=(x_t-x_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差, 注意到 $\sum\limits_{i=1}^{m} l_i=1$, 则有

$$e_t=x_t-\hat{x}_t=x_t-\sum_{i=1}^{m} l_i x_{it}=\sum_{i=1}^{m} l_i(x_t-x_{it})=\sum_{i=1}^{m} l_i e_{it},\quad t=1,2,\cdots,n.$$

常见的衡量组合预测精度的一个准则就是使组合预测误差平方和达到最小. 实际上绝对误差是衡量预测精度的另一个指标, 它与预测误差平方和相比的一个显著特点是有比较好的稳健性. 为此先构造基于 OWA 算子的组合预测绝对误差.

根据定义 8.1.1, 令

$$\mathrm{OWA}_W(|e_1|,|e_2|,\cdots,|e_n|)=\sum_{i=1}^{n} w_i\,|d_i|. \tag{8.5.1}$$

则式 (8.5.1) 称为由绝对误差序列 $|e_1|,|e_2|,\cdots,|e_n|$ 所产生的基于 OWA 算子的组合预测绝对误差. 其中 $|d_i|$ 是 $|e_1|,|e_2|,\cdots,|e_n|$ 中按从大到小的顺序排列的第 i 个大的数, $i=1,2,\cdots,n$, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 是与 f_W 有关的加权向量, 可根据实际情况进行选取. 特别地, 当 $\boldsymbol{W}=\left(\dfrac{1}{n},\dfrac{1}{n},\cdots,\dfrac{1}{n}\right)^{\mathrm{T}}$ 时, 基于 OWA 算子的组合预测绝对误差就是常见的组合预测的平均绝对误差.

为了使组合预测值逼近实际值, 我们希望 OWA 算子的组合预测绝对误差愈小愈好. 也就是 $\mathrm{OWA}_W(|e_1|,|e_2|,\cdots,|e_n|)$ 越小表示组合预测方法越有效. 因此, 基于

OWA 算子的组合预测模型可表示为如下形式

$$\min \mathrm{OWA}_W(|e_1|,|e_2|,\cdots,|e_n|)=\sum_{i=1}^{n} w_i|d_i|,$$

$$\text{s.t.}\begin{cases} e_t=\sum_{i=1}^{m} l_i e_{it}, \quad t=1,2,\cdots,n, \\ \sum_{i=1}^{m} l_i=1, \quad l_1,l_2,\cdots,l_m \geqslant 0. \end{cases} \tag{8.5.2}$$

在预测过程中, 有时预测者可能关心最大的绝对误差. 因此, 当 OWA 算子的加权向量 $\boldsymbol{W}=(1,0,\cdots,0)^{\mathrm{T}}$ 时, 基于 OWA 算子的组合预测绝对误差为最大绝对误差. 即

$$\mathrm{OWA}_{(1,0,\cdots,0)}(|e_1|,|e_2|,\cdots,|e_n|)=\max_{1\leqslant t\leqslant n}|e_t|. \tag{8.5.3}$$

可见组合预测的最大绝对误差就是基于 OWA 算子的组合预测绝对误差的特例.

显然, 组合预测的最大绝对误差 $\max\limits_{1\leqslant t\leqslant n}|e_t|$ 为各种单项预测方法的加权系数 $l_1,l_2,\cdots,l_m$ 的函数, 记为 $E(l_1,l_2,\cdots,l_m)$. 即

$$E(l_1,l_2,\cdots,l_m)=\max_{1\leqslant t\leqslant n}|e_t|. \tag{8.5.4}$$

因此, 基于最大绝对误差最小化的组合预测模型 (8.5.2) 可表示为如下形式

$$\min E(l_1,l_2,\cdots,l_m)=\min\max_{1\leqslant t\leqslant n}|e_t|,$$

$$\text{s.t.}\begin{cases} e_t=\sum_{i=1}^{m} l_i e_{it}, \quad t=1,2,\cdots,n, \\ \sum_{i=1}^{m} l_i=1, \quad l_1,l_2,\cdots,l_m \geqslant 0. \end{cases} \tag{8.5.5}$$

组合预测模型 (8.5.5) 可化为如下线性规划问题求得最优解. 令 $v=\max\limits_{1\leqslant t\leqslant n}|e_t|$, 所以

$$|e_t|\leqslant v, \quad t=1,2,\cdots,n.$$

则有

$$\min v,$$

$$\text{s.t.}\begin{cases} \sum_{i=1}^{m} l_i e_{it}-v\leqslant 0, \quad t=1,2,\cdots,n, \\ \sum_{i=1}^{m} l_i e_{it}+v\geqslant 0, \quad t=1,2,\cdots,n, \\ \sum_{i=1}^{m} l_i=1, \quad l_1,l_2,\cdots,l_m \geqslant 0. \end{cases} \tag{8.5.6}$$

这是一个线性规划问题, 可有现成的 Matlab 最优化工具箱来求解.

8.5.2 一类基于 OWA 算子的组合预测模型的性质[151]

在下面的讨论中，我们针对组合预测的最大误差绝对值达到最小的这样特殊一类基于 OWA 算子的组合预测模型研究它的非劣性组合预测、冗余预测方法的存在性以及冗余信息的判定定理.

1. 几个概念

令 $E_i=\max\{|e_{it}|,t=1,2,\cdots,n\}$, 则 E_i 表示第 i 种单项预测方法的最大绝对误差, $E_{\min}$ 表示 m 种单项预测方法的最大绝对误差中的最小者, $E_{\max}$ 表示 m 种单项预测方法的最大绝对误差中的最大者, 即

$$E_{\min}=\min_{1\leqslant i\leqslant m}E_i,\quad E_{\max}=\max_{1\leqslant i\leqslant m}E_i.$$

定义 8.5.1 设 $E(l_1,l_2,\cdots,l_m)$ 为组合预测的最大绝对误差, 若 $E(l_1,l_2,\cdots,l_m)<E_{\min}$, 则称权系数 $l_1,l_2,\cdots,l_m$ 确定的组合预测模型为优性组合预测, 若 $E_{\min}\leqslant E(l_1,l_2,\cdots,l_m)\leqslant E_{\max}$, 则称之为非劣性组合预测, 若 $E(l_1,l_2,\cdots,l_m)>E_{\max}$, 则称之为劣性组合预测.

定义 8.5.1 表明, 只有组合预测的最大绝对误差小于各单项预测方法的最大绝对误差中的最小者, 该组合预测模型才是优性的. 即以最大绝对误差最小化作为判断准则, 优性组合预测一定比 “最好” 的单项预测方法还要 “好”.

定义 8.5.2 若第 i 种和第 k 种单项预测方法在各个时刻预测误差的绝对值满足如下不等式 $|e_{it}|\leqslant|e_{kt}|,t=1,2,\cdots,n$, 且至少对某个时刻 t_0 有严格的不等号成立, $t_0\in\{1,2,\cdots,n\}$, 则称第 i 种单项预测方法优超第 k 种单项预测方法.

定义 8.5.2 表明, 第 i 种单项预测方法在各个时刻预测误差的绝对值不大于第 k 种单项预测方法, 而且至少对某个时刻 t_0 处预测误差的绝对值严格小于第 k 种单项预测方法, 直观上可以认为第 i 种单项预测方法要 “好于” 第 k 种单项预测方法.

2. 非劣性组合预测和冗余预测方法的判定

定理 8.5.1 组合预测的最大绝对误差不超过各单项预测方法最大绝对误差的加权平均值. 即

$$E(l_1,l_2,\cdots,l_m)\leqslant\sum_{i=1}^{m}l_iE_i. \tag{8.5.7}$$

证明 设 $\boldsymbol{L}=(l_1,l_2,\cdots,l_m)^{\mathrm{T}}$ 为组合预测模型 (8.5.5) 的任一个可行解, 不妨设在 t_0 时刻 ($t_0\in\{1,2,\cdots,n\}$) 组合预测的绝对误差为最大值, 即 $|e_{t_0}|=\max\limits_{1\leqslant t\leqslant n}|e_t|$, 因为由 E_i 的定义知 $E_i=\max\{|e_{it}|,t=1,2,\cdots,n\}$, 所以 $|e_{it_0}|\leqslant E_i$, 并注意到式

(8.5.1), 从而有

$$E(l_1,l_2,\cdots,l_m)=\max_{1\leqslant t\leqslant n}|e_t|=|e_{t_0}|=\left|\sum_{i=1}^{m}l_ie_{it_0}\right|\leqslant\sum_{i=1}^{m}l_i|e_{it_0}|\leqslant\sum_{i=1}^{m}l_iE_i.$$ 证毕.

推论 8.5.1 组合预测模型 (8.5.5) 的任一可行解对应的组合预测至少是非劣性组合预测.

推论 8.5.2 简单加权平均组合预测方法至少是非劣性组合预测.

定理 8.5.2 组合预测模型 (8.5.5) 的最优目标函数值是参与组合预测的各单项预测方法总个数 m 的单调不增函数, 即

$$E(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})\leqslant E(l_1^*,l_2^*,\cdots,l_m^*),\tag{8.5.8}$$

其中 $E(l_1^*,l_2^*,\cdots,l_m^*)$ 和 $E(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})$ 分别表示 m 个单项预测方法及再增加一个单项预测方法共 $m+1$ 个单项预测方法参与的基于 OWA 算子的组合预测模型对应的最优目标函数值.

证明 设对 m 个单项预测方法参与的组合预测模型 (8.5.5) 的任一个可行解, 不妨设在 t_0 时刻 ($t_0\in\{1,2,\cdots,n\}$) 组合预测的绝对误差为最大值, 即 $|e_{t_0}|=\max\limits_{1\leqslant t\leqslant n}|e_t|=\left|\sum\limits_{i=1}^{m}l_ie_{it_0}\right|$.

设 $\boldsymbol{L}^*=(l_1^*,l_2^*,\cdots,l_m^*)^{\mathrm{T}}$ 为 m 个单项预测方法参与的组合预测模型 (8.5.5) 的最优解, 其最优目标函数值为

$$E(l_1^*,l_2^*,\cdots,l_m^*)=\min_{l_1,l_2,\cdots,l_m,\sum\limits_{i=1}^{m}l_i=1}\max_{1\leqslant t\leqslant n}|e_t|=\min_{l_1,l_2,\cdots,l_m,\sum\limits_{i=1}^{m}l_i=1}|e_{t_0}|=\left|\sum_{i=1}^{m}l_i^*e_{it_0}\right|,\tag{8.5.9}$$

其中 $\sum\limits_{i=1}^{m}l_i^*=1,\ l_i^*\geqslant 0,\ i=1,2,\cdots,m.$

同理, 再增加一个单项预测方法共 $m+1$ 个单项预测方法参与的组合预测模型 (8.5.5) 的任一个可行解, 不妨设在 t_1 时刻 ($t_1\in\{1,2,\cdots,n\}$) 组合预测的绝对误差为最大值, $|e_{t_1}|=\max\limits_{1\leqslant t\leqslant n}|e_t|=\left|\sum\limits_{i=1}^{m+1}l_ie_{it_1}\right|$, 设 $\bar{\boldsymbol{L}}=(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1})^{\mathrm{T}}$ 为再增加一个单项预测方法, 共 $m+1$ 个单项预测方法参与的组合预测模型 (8.5.5) 的最优解, 其最优目标函数值为

$$E(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1}) = \min_{l_1,l_2,\cdots,l_m,l_{m+1},\sum\limits_{i=1}^{m+1} l_i=1} \max_{1\leqslant t\leqslant n} |e_t|$$
$$= \min_{l_1,l_2,\cdots,l_m,l_{m+1},\sum\limits_{i=1}^{m+1} l_i=1} |e_{t_1}| = \left|\sum_{i=1}^{m+1} \bar{l}_i e_{it_1}\right|, \quad (8.5.10)$$

其中 $\sum\limits_{i=1}^{m+1} \bar{l}_i = 1, \bar{l}_i \geqslant 0, i=1,2,\cdots,m+1.$

令 $\boldsymbol{L}=(l_1^*,l_2^*,\cdots,l_m^*,0)^{\mathrm{T}}$, 显然, $\boldsymbol{L}$ 为 m+1 个单项预测方法参与的组合预测模型 (8.5.5) 的可行解, 则有

$$E(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1}) \leqslant E(l_1^*,l_2^*,\cdots,l_m^*,0). \quad (8.5.11)$$

由式 (8.5.10) 得

$$E(l_1^*,l_2^*,\cdots,l_m^*,0) = \left|\sum_{i=1}^{m} l_i^* e_{it_1}\right|. \quad (8.5.12)$$

因为 $\max\limits_{1\leqslant t\leqslant n}|e_t| = |e_{t_0}|$, 所以 $|e_{t_1}| \leqslant |e_{t_0}|$, 注意到式 $e_t=\sum\limits_{i=1}^{m} l_i e_{it}$, 则有

$$\left|\sum_{i=1}^{m} l_i^* e_{it_1}\right| \leqslant \left|\sum_{i=1}^{m} l_i^* e_{it_0}\right|. \quad (8.5.13)$$

由式 (8.5.9)、(8.5.11)~(8.5.13) 即得

$$E(\bar{l}_1,\bar{l}_2,\cdots,\bar{l}_m,\bar{l}_{m+1}) \leqslant E(l_1^*,l_2^*,\cdots,l_m^*).$$ 证毕.

通常认为随着参与组合预测的单项预测方法个数 m 的增加, 基于 OWA 算子的组合预测模型的最优目标函数值一定是 m 的严格单调减少函数, 然而, 定理 8.5.2 证明了当再增加一个单项预测方法时, 基于 OWA 的组合预测模型的最优目标函数值可能减少, 也可能不变. 这表明组合预测模型可能存在冗余的单项预测方法. 下面一个定理为冗余预测方法提供了判定.

定理 8.5.3 若对任意第 t 时刻 $e_{1t},e_{2t},\cdots,e_{mt}$ 的符号完全相同, $t=1,2,\cdots,n$, 且第 i 种单项预测方法优超第 k 种单项预测方法, 则组合预测模型 (8.5.5) 的冗余度至少为 $1/m$.

证明 假设 $\boldsymbol{L}^*=(l_1^*,\cdots,l_i^*,\cdots,l_k^*,\cdots,l_m^*)^{\mathrm{T}}$ 为组合预测模型 (8.5.5) 的最优解且第 k 种单项预测方法不为冗余预测方法, 即 $l_k^*>0$, 且 $\sum\limits_{i=1}^{m} l_i^*=1$, 由式 (8.5.9) 知最优解 $\boldsymbol{L}^*$ 对应的组合预测模型 (8.5.5) 的目标函数值为

$$E(l_1^*,l_2^*,\cdots,l_m^*) = \left|\sum_{i=1}^{m} l_i^* e_{it_0}\right|. \quad (8.5.14)$$

因为对任意第 t 时刻 $e_{1t}, e_{2t}, \cdots, e_{mt}$ 的符号完全相同, $t \in \{1, 2, \cdots, n\}$, 所以 $\left|\sum_{i=1}^{m} l_i^* e_{it_0}\right| = \sum_{i=1}^{m} l_i^* |e_{it_0}|$. 则

$$E(l_1^*, l_2^*, \cdots, l_m^*) = \sum_{i=1}^{m} l_i^* |e_{it_0}|. \tag{8.5.15}$$

又因为第 i 种单项预测方法优超第 k 种单项预测方法, 由定义 2 知 $|e_{it}| \leqslant |e_{kt}|$, $t = 1, 2, \cdots, n$, 且至少对某个 t_0 有严格的不等号成立, $t_0 \in \{1, 2, \cdots, n\}$, 所以有

$$|e_{it_0}| < |e_{kt_0}|, \quad t = 1, 2, \cdots, n. \tag{8.5.16}$$

构造 $\tilde{\boldsymbol{L}} = (l_1^*, \cdots, l_i^* + l_k^*, \cdots, 0, \cdots, l_m^*)^{\mathrm{T}}$, $\tilde{\boldsymbol{L}}$ 的第 i 个分量为 $l_i^* + l_k^*$, 第 k 个分量为 0, 其余 $m-2$ 个分量与最优解 $\boldsymbol{L}^*$ 相同, 显然 $\tilde{\boldsymbol{L}}$ 为组合预测模型 (8.5.5) 的一个可行解, 则与 $\tilde{\boldsymbol{L}}$ 对应组合预测模型 (8.5.5) 的目标函数值为

$$\begin{aligned} E(\tilde{\boldsymbol{L}}) &= |l_1^* e_{1t_0} + l_2^* e_{2t_0} + \cdots + (l_i^* + l_k^*) e_{it_0} + \cdots + 0 e_{kt_0} + \cdots + l_n^* e_{nt_0}| \\ &= l_1^* |e_{1t_0}| + l_2^* |e_{2t_0}| + \cdots + (l_i^* + l_k^*) |e_{it_0}| + \cdots + 0 |e_{kt_0}| + \cdots + l_n^* |e_{nt_0}|. \end{aligned} \tag{8.5.17}$$

由式 (8.5.15)~(8.5.17) 知

$$E(\tilde{\boldsymbol{L}}) < \sum_{i=1}^{m} l_i^* |e_{it_0}| = E(l_1^*, l_2^*, \cdots, l_m^*).$$

而这与 $\boldsymbol{L}^* = (l_1^*, \cdots, l_i^*, \cdots, l_k^*, \cdots, l_m^*)^{\mathrm{T}}$ 为组合预测模型 (8.5.5) 的最优解矛盾! 所以假设不成立. 从而第 k 种单项预测方法一定是冗余预测方法, 此即组合预测模型 (8.5.5) 的冗余度至少为 $1/m$. 证毕.

本节提出了基于有序加权平均 (OWA) 算子的组合预测模型, 针对最大绝对误差最小化的组合预测模型的性质作了研究, 得出一些有益的结果, 在理论上进一步表明该类最优组合预测方法确实能综合各种单项预测方法信息. 从而为该模型的广泛应用提供理论依据.

第 9 章　组合预测技术的应用研究

为了综合利用各种预测方法所提供的信息, 在预测实践中, 采用组合预测的办法. 即对不同的预测模型以适当的加权平均形式组合起来, 得出组合预测模型. 组合预测核心的问题就是在一定的准则下求出最优加权平均系数. 这种最优加权综合的思想有着广泛的应用, 本章主要讨论组合预测技术在决策、评价、证券投资、劳动力配置等方面的应用.

9.1　多属性决策中最优组合赋权方法研究

9.1.1　组合赋权方法概述 [115]

多属性决策是多目标决策的一种, 是指对于给定的有限个选择方案, 决策者根据事先确定的各个方案若干个属性值, 按照某种决策准则进行多方案排序. 多属性决策广泛应用于社会、经济、管理等诸多领域. 对于多属性决策问题, 无论采取什么样的求解方法, 一般需要确定各属性的相对重要程度, 而重要程度往往用属性的权系数来反映, 权系数越大则其对应的属性就越重要. 因此, 权系数的正确确定, 对于多属性决策问题的正确决策具有十分重要的作用.

目前权系数的确定方法有多种. 大体上可分为主观赋权方法和客观赋权方法两大类 [112]. 主观赋权法是决策者根据经验主观判断或各指标的主观重视程度进行赋权的方法. 如专家调查法、二项系数法、AHP 法等; 而客观赋权法是通过建立一定的数学模型计算出权重系数, 如主成分分析法、熵技术法、均方差法及目标规划法等. 两大类赋权方法各有不同的特点. 主观赋权法的随意性较大, 决策准确性和可靠性稍差一些, 这是其不足之处. 但属性的相对重要程度一般不会违反人们的常识. 客观赋权法显著的特点是存在赋权的客观标准, 通过计算得出属性的权重系数, 而不是人为给定的. 但客观赋权法的缺点是有时计算结果无法解释.

从上述主观赋权法和客观赋权法的特点分析可知, 它们均具有一定的互补性. 为了让多属性决策的排序结果更科学, 一个合理的做法就是将不同的赋权法所得的权重系数按照一定的方法进行组合. 组合赋权是一种新的权重系数确定方法, 根据多属性决策问题的实际情况, 选择几种有代表性赋权的方法, 通过适当的数学模型, 求组合赋权系数. 所谓的代表性, 是指在进行组合赋权时应该在主观赋权法和客观赋权法中至少各选出一种方法. 这样通过组合赋权, 可以充分利用各种赋权方法所带来的信息, 使得排序结果既能体现主观信息, 又能体现客观信息.

文献 [112] 指出将主、客观信息综合集成的方法是一个有价值的新课题. 文献 [112] 提出了一种主客观赋权法, 该方法以各决策方案的评价目标值之和达到最大为目标函数, 建立一个数学规划模型来求解属性权重的. 文献 [113] 又提出了另外一种权重集成方法, 即设 G 为组合权重与各决策者给出的主观权重之间的总偏差平方和, 设 G' 为各选择方案与其 "理想方案" 之间评价目标值的偏差平方和, 以 G 和 G' 的加权平均达到最小为目标函数, 建立一个规划模型来求解组合属性权重. 文献 [114] 提出了一种确定多指标权系数的离差平方和最大化方法. 目前优化组合赋权方法的研究还不完善, 有必要进一步探讨多种准则下的优化组合赋权方法, 并进行比较分析. 本节在文献 [114] 的基础上提出了多属性决策中基于离差平方和的新的最优组合赋权方法, 利用该方法对文献 [112,113] 中的实例进行了计算, 结果与文献 [112,113] 是一致的. 这就表明该最优组合赋权方法的有效性和科学合理性.

9.1.2 基于离差平方和的最优组合赋权方法的基本原理

设有某个多属性决策问题, 其方案集表示为 $S=\{S_1,S_2,\cdots,S_m\}$, 其属性 (或指标) 集表示为 $P=\{P_1,P_2,\cdots,P_n\}$, 第 i 个方案 S_i 对第 j 个属性 P_j 的属性值记为 a_{ij} , $i=1,2,\cdots,m, j=1,2,\cdots,n$, 矩阵 $\boldsymbol{A}=\left(a_{ij}\right)_{m\times n}$ 称为属性矩阵或决策矩阵. 通常 属性可分为效益型、成本型、固定型、区间型.

所谓效益型属性是指属性值愈大愈好的指标, 如资金产值率、资金利税率、全员劳动生产率等. 所谓成本型属性是指属性值愈小愈好的指标, 如流动资金占用额、流动资金周转天数等. 所谓固定型属性是指属性值既不能过大又不能过小, 而以稳定在某个固定值为最佳的一类指标, 家用电器稳压器的稳压性能指标就属于这类指标. 所谓区间型属性是指属性值以落在某个固定区间内为最佳的一类指标, 如财务评价中的流动比率指标也可以看成是这类指标, 流动比率越高, 则资产的流动性较大, 短期偿债能力越强. 然而, 过高的流动比率可能意味着企业运用资金的效率不高. 因此, 流动比率应存在一个最合适的区间. 由于不同的属性往往具有不同的量纲和量纲单位, 为了消除它们带来的不可公度性, 在决策之前首先应将属性指标作无量纲化处理. 然而决策属性类型不同, 无量纲化处理方法也将不同.

对于效益型属性, 一般可令

$$b_{ij}=\frac{a_{ij}-a_j^{\min}}{a_j^{\max}-a_j^{\min}}. \tag{9.1.1}$$

对于成本型属性, 一般可令

$$b_{ij}=\frac{a_j^{\max}-a_{ij}}{a_j^{\max}-a_j^{\min}}. \tag{9.1.2}$$

对于固定型属性, 一般可令

$$b_{ij} = 1 - \frac{|a_{ij} - a_j^*|}{\max\limits_i |a_{ij} - a_j^*|}, \tag{9.1.3}$$

其中 a_j^* 为第 j 个属性 P_j 的最佳稳定值.

对于区间型属性, 一般可令

$$b_{ij} = \begin{cases} 1 - \dfrac{q_{1j} - a_{ij}}{\max\limits_i\{q_{1j} - a_j^{\min}, a_j^{\max} - q_{2j}\}}, & 当a_{ij} < q_{1j}, \\ 1, & 当a_{ij} \in [q_{1j}, q_{2j}], \\ 1 - \dfrac{a_{ij} - q_{2j}}{\max\limits_i\{q_{1j} - a_j^{\min}, a_j^{\max} - q_{2j}\}}, & 当a_{ij} > q_{2j}, \end{cases} \tag{9.1.4}$$

式中 $a_j^{\min}, a_j^{\max}$ 分别为第 j 个属性 P_j 的最小值和最大值. $[q_{1j}, q_{2j}]$ 为 P_j 的最佳稳定区间.

显然, $b_{ij} \in [0,1], i = 1,2,\cdots,m, j = 1,2,\cdots,n$. 称上述的无量纲化处理的矩阵 $\boldsymbol{B} = (b_{ij})_{m\times n}$ 为规范化的决策矩阵, b_{ij} 表示第 i 个方案 S_i 对第 j 个属性 P_j 的规范化属性值, 矩阵 $\boldsymbol{B}$ 的第 i 行表示第 i 个方案 S_i 对 n 个属性值的规范值. 显然, b_{ij} 愈大愈好.

不管是主观赋权方法还是客观赋权方法, 前已指出它们均有若干具体的赋权方法. 假设某个多属性决策问题, 对 n 个属性有 l 种具体的赋权方法对其赋值. 设第 k 种赋权方法给出的权向量值为

$$\boldsymbol{W}_k = (w_{1k}, w_{2k}, \cdots, w_{nk})^{\mathrm{T}}, \quad k = 1,2,\cdots,l,$$

其中 $w_{jk} \geqslant 0, \sum\limits_{j=1}^{n} w_{jk} = 1, \quad k = 1,2,\cdots,l, j = 1,2,\cdots,n.$

为综合各种赋权方法的特点, 可考虑如下组合赋权$\boldsymbol{W}_c=(w_{c1}, w_{c2}, \cdots, w_{cn})^{\mathrm{T}}$, 令

$$\boldsymbol{W}_c = \theta_1\boldsymbol{W}_1 + \theta_2\boldsymbol{W}_2 + \theta_l\mathbf{W}_l, \tag{9.1.5}$$

称 $\boldsymbol{W}_c = (w_{c1}, w_{c2}, \cdots, w_{cn})^{\mathrm{T}}$ 为组合赋权系数向量. 其中 $\theta_1, \theta_2, \cdots, \theta_l$ 为组合权系数向量的线性表出系数. $\theta_k \geqslant 0, k = 1,2,\cdots,l$, 且满足单位化约束条件

$$\sum_{k=1}^{l} \theta_k^2 = 1. \tag{9.1.6}$$

令分块矩阵 $\boldsymbol{W} = (\boldsymbol{W}_1, \boldsymbol{W}_2, \cdots, \boldsymbol{W}_l)$, $\boldsymbol{\Theta} = (\theta_1, \theta_2, \cdots, \theta_l)^{\mathrm{T}}$, 则称 $\boldsymbol{W}$ 为 l 个权系数向量组成的矩阵. $\boldsymbol{W}$ 实际上为 $n \times l$ 的矩阵, $\boldsymbol{\Theta}$ 为组合权系数向量的线性表出

系数组成的 l 维列向量. 此时式 (9.1.5)、(9.1.6) 可表为矩阵形式

$$\boldsymbol{W}_c = \boldsymbol{W}\boldsymbol{\Theta}, \tag{9.1.7}$$

$$\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta} = 1. \tag{9.1.8}$$

根据简单线性加权法, 由组合赋权系数向量 $\boldsymbol{W}_c$ 计算而得的第 i 个决策方案 S_i 的多属性综合评价值可表示为

$$D_i(\boldsymbol{W}_c) = \sum_{j=1}^{n} b_{ij} w_{cj}, \quad i = 1, 2, \cdots, m. \tag{9.1.9}$$

一般而言, $D_i(\boldsymbol{W}_c)$ 总是愈大愈好, $D_i(\boldsymbol{W}_c)$ 愈大表示第 i 个决策方案 S_i 愈优. 但是, 在多属性决策中, 如果各属性的权系数确定不当, 致使各决策方案的多属性综合评价值 $D_i(\boldsymbol{W}_c)(i = 1, 2, \cdots, m)$ 互相差别很小, 这样将不利于决策方案的排序. 所以选择组合赋权系数向量的一个基本思想是使各决策方案的多属性综合评价值 $D_i(\boldsymbol{W}_c)$ 尽可能分散, 越分散越有利于方案的决策与排序 [114]. 本节采用文献 [114] 中把各决策方案多属性综合评价值的离差平方和作为其分散程度的度量. 为此, 定义第 i_1 个决策方案综合评价值 $D_{i_1}(\boldsymbol{W}_c)$ 和第 i_2 个决策方案综合评价值 $D_{i_2}(\boldsymbol{W}_c)$ 的离差为

$$v_{i_1 i_2}(\boldsymbol{W}_c) = D_{i_1}(\boldsymbol{W}_c) - D_{i_2}(\boldsymbol{W}_c) = \sum_{j=1}^{n} (b_{i_1 j} - b_{i_2 j}) w_{cj}, \quad i_1, i_2 = 1, 2, \cdots, m. \tag{9.1.10}$$

设 $v_i(\boldsymbol{W}_c)$ 表示第 i 个决策方案与其他各决策方案综合评价值的离差平方和, 则有

$$v_i(\boldsymbol{W}_c) = \sum_{i_1=1}^{m} v_{ii_1}^2(\boldsymbol{W}_c) = \sum_{i_1=1}^{m} \left[\sum_{j=1}^{n} (b_{ij} - b_{i_1 j}) w_{cj}\right]^2, \quad i = 1, 2, \cdots, m. \tag{9.1.11}$$

根据前述的选择组合赋权系数向量 $\boldsymbol{W}_c$ 的基本思想, 应该使 m 个决策方案总的离差平方和达到最大, 这样有利于方案的决策与排序. 于是可构造如下目标函数 [115]

$$\begin{aligned} J_1(\boldsymbol{W}_c) &= \sum_{i=1}^{m} v_i(\boldsymbol{W}_c) = \sum_{i=1}^{m}\sum_{i_1=1}^{m} \left[\sum_{j=1}^{n} (b_{ij} - b_{i_1 j}) w_{cj}\right]^2 \\ &= \sum_{i=1}^{m}\sum_{i_1=1}^{m} \left[\sum_{j_1=1}^{n}\sum_{j_2=1}^{n} (b_{ij_1} - b_{i_1 j_1}) w_{cj_1} (b_{ij_2} - b_{i_1 j_2}) w_{cj_2}\right] \\ &= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} \left[\sum_{i=1}^{m}\sum_{i_1=1}^{m} (b_{ij_1} - b_{i_1 j_1})(b_{ij_2} - b_{i_1 j_2})\right] w_{cj_1} w_{cj_2}. \end{aligned} \tag{9.1.12}$$

若令 $n\times n$ 阶矩阵 $\boldsymbol{B}_1$ 为

$$\boldsymbol{B}_1=\begin{bmatrix} a_1^* & a_2^* & \cdots & a_n^* \\ b_1^* & b_2^* & \cdots & b_n^* \\ c_1^* & c_2^* & \cdots & c_n^* \end{bmatrix},$$

其中,

$$a_1^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{i1}-b_{i_11})(b_{i1}-b_{i_11}),$$

$$a_2^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{i1}-b_{i_11})(b_{i2}-b_{i_12}),$$

$$\cdots\cdots\cdots\cdots$$

$$a_n^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{i1}-b_{i_11})(b_{in}-b_{i_1n}),$$

$$b_1^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{i2}-b_{i_12})(b_{i1}-b_{i_11}),$$

$$b_2^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{i2}-b_{i_12})(b_{i2}-b_{i_12}),$$

$$\cdots\cdots\cdots\cdots$$

$$b_n^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{i2}-b_{i_12})(b_{in}-b_{i_1n}),$$

$$c_1^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{in}-b_{i_1n})(b_{i1}-b_{i_11}),$$

$$c_2^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{in}-b_{i_1n})(b_{i2}-b_{i_12}),$$

$$\cdots\cdots\cdots\cdots$$

$$c_n^*=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{in}-b_{i_1n})(b_{in}-b_{i_1n}).$$

则目标函数 $J_1(\boldsymbol{W}_c)$ 可表为

$$J_1(\boldsymbol{W}_c)=\boldsymbol{W}_c^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}_c, \tag{9.1.13}$$

其中 $\mathbf{W}_c=(w_{c1},w_{c2},\cdots,w_{cn})^{\mathrm{T}}$ 为组合赋权系数向量.

对于 n 阶对称方阵 $\boldsymbol{B}_1$, 我们有如下引理.

引理 9.1.1 $\boldsymbol{B}_1$ 为 n 阶对称方阵, 且 $\boldsymbol{B}_1$ 为非负定矩阵.

证明 因为 $\boldsymbol{B}_1$ 的第 s 行第 t 列的元素为

$$\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{is}-b_{i_1s})(b_{it}-b_{i_1t}),$$

$\boldsymbol{B}_1$ 的第 t 行第 s 列的元素为

$$\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{it}-b_{i_1t})(b_{is}-b_{i_1s}),$$

显然有

$$\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{is}-b_{i_1s})(b_{it}-b_{i_1t})=\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{it}-b_{i_1t})(b_{is}-b_{i_1s}).$$

从而 $\boldsymbol{B}_1$ 为 n 阶对称方阵. 对任意的 n 维向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_l)^{\mathrm{T}}\neq 0$, 则有

$$\begin{aligned}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{x}&=\sum_{s=1}^{n}\sum_{t=1}^{n}\left[\sum_{i=1}^{m}\sum_{i_1=1}^{m}(b_{is}-b_{i_1s})(b_{it}-b_{i_1t})\right]x_sx_t\\&=\sum_{i=1}^{m}\sum_{i_1=1}^{m}\left[\sum_{s=1}^{n}(b_{ij}-b_{i_1j})x_s\right]^2\geqslant 0.\end{aligned}$$

从而 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{x}$ 为非负定二次型, 所以 $\boldsymbol{B}_1$ 为非负定矩阵. 证毕.

要想求出组合赋权系数向量 $\boldsymbol{W}_c$, 由式 (9.1.7) 知, 只需求出组合权系数的线性表出系数向量 $\boldsymbol{\Theta}$. 既然 $\boldsymbol{W}_c$ 为 $\boldsymbol{\Theta}$ 的函数, 从而式 (9.1.13) 表明目标函数 $J_1(\boldsymbol{W}_c)$ 也为 $\boldsymbol{\Theta}$ 的函数, 此时, $J_1(\boldsymbol{W}_c)$ 可记为 $F(\boldsymbol{\Theta})$. 于是, 基于 m 个决策方案总的离差平方和的最优组合赋权方法即为如下最优化问题[115]

$$\begin{aligned}&\max F(\boldsymbol{\Theta})=\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}\boldsymbol{\Theta},\\&\text{s.t.}\begin{cases}\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta}=1,\\\boldsymbol{\Theta}\geqslant 0.\end{cases}\end{aligned}\tag{9.1.14}$$

对于模型 (9.1.14), 先不考虑 $\boldsymbol{\Theta}$ 的非负性, 最优化问题 (9.1.14) 可简化为无约束优化问题

$$\max F_1(\boldsymbol{\Theta})=\frac{\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}\boldsymbol{\Theta}}{\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta}}.\tag{9.1.15}$$

对于式 (9.1.15) , 我们有如下结论.

定理 9.1.1 $\max F_1(\boldsymbol{\Theta})=\dfrac{\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}\boldsymbol{\Theta}}{\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta}}=\lambda_{\max}$, 最优解 $\boldsymbol{\Theta}^*$ 为 $\lambda_{\max}$ 对应的单位化特征向量. 其中 $\lambda_{\max}$ 为矩阵 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 的最大特征根.

证明 记 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\mathbf{W}=\boldsymbol{G}$, 因为 $\boldsymbol{B}_1$ 为 n 阶对称方阵, 显然有 $\boldsymbol{G}^{\mathrm{T}}=\boldsymbol{G}$, 所以 $\boldsymbol{G}$ 也是 l 阶的对称方阵. 设 $\lambda_1,\lambda_2,\cdots,\lambda_l$ 为 $\boldsymbol{G}$ 的特征值, 不妨假定 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_l$, 即 $\lambda_1=\lambda_{\max}$.

记 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_l$ 分别为 $\lambda_1,\lambda_2,\cdots,\lambda_l$ 的对应的标准正交化的特征向量, 即

$$\boldsymbol{G}\boldsymbol{p}_i=\lambda_i\boldsymbol{p}_i,\quad i=1,2,\cdots,l \tag{9.1.16}$$

令 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_l)$, $\boldsymbol{\Lambda}=\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_l)$, 即 $\boldsymbol{\Lambda}$ 为 $\boldsymbol{G}$ 的特征值所构成的对角阵, 由式 (9.1.16), 则有

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}=\boldsymbol{\Lambda}. \tag{9.1.17}$$

因为 $\boldsymbol{P}$ 为正交矩阵, 因而是可逆的. 对任意的向量 $\boldsymbol{\Theta}$, 方程组

$$\boldsymbol{P}\boldsymbol{x}=\boldsymbol{\Theta} \tag{9.1.18}$$

总存在唯一解. 设其解 $\boldsymbol{x}=(x_1,x_2,\cdots,x_l)^{\mathrm{T}}$, 所以

$$\begin{aligned}F_1(\boldsymbol{\Theta})&=\frac{(\boldsymbol{P}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x}}{(\boldsymbol{P}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}}=\frac{(\boldsymbol{P}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x}}{(\boldsymbol{P}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}}=\frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}}\\&=\frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}=\sum_{i=1}^{l}\omega_i\lambda_i,\end{aligned} \tag{9.1.19}$$

其中 $\omega_i=\dfrac{x_i^2}{\sum\limits_{i=1}^{l}x_i^2}\geqslant 0$, 注意到 $\sum\limits_{i=1}^{l}\omega_i=1$, 则有

$$F_1(\boldsymbol{\Theta})=\sum_{i=1}^{l}\omega_i\lambda_i\leqslant\lambda_1\sum_{i=1}^{l}\omega_i=\lambda_{\max}. \tag{9.1.20}$$

式 (9.1.20) 表明等式 $F_1(\boldsymbol{\Theta}^*)=\lambda_{\max}$ 成立的充要条件是

$$\omega_1=1,\quad\omega_i=0,\quad i=2,\cdots,l. \tag{9.1.21}$$

由式 (9.1.18)、(9.1.21) 得 $\boldsymbol{\Theta}^*=\boldsymbol{p}_1$. 证毕.

定理 9.1.1 表明, $F_1(\boldsymbol{\Theta})$ 最大值为对称矩阵 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 最大特征根, 最优解 $\boldsymbol{\Theta}^*$ 为矩阵 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 的最大特征根所对应的单位化特征向量.

由引理 9.1.1 知, 矩阵 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 是对称非负定的, 根据非负不可约矩阵的 Perron-Frobenius 定理, $\lambda_{\max}$ 为单根, 且它对应的 $\boldsymbol{\Theta}^*$ 的分量全部为正. 因此, $\boldsymbol{\Theta}^*$ 也是模型 (9.1.14) 的最优解. 求解 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 的最大特征根所对应的单位化特征向

量 $\boldsymbol{\Theta}^*$, 可以采用现成的计算软件, 如 Matlab, 当然也可采用幂法迭代算法 [114], 步骤如下:

(1) 给定初始的单位化向量 $\boldsymbol{\Theta}(0)=(\theta_1(0),\theta_2(0),\cdots,\theta_l(0))^{\mathrm{T}}>0$, 一般可取

$$\boldsymbol{\Theta}(0)=\left(\frac{1}{\sqrt{l}},\frac{1}{\sqrt{l}},\cdots,\frac{1}{\sqrt{l}}\right)^{\mathrm{T}},$$

同时给出迭代精度 $\varepsilon=10^{-6}$, 置 $k=0$.

(2) 计算 $\tilde{\boldsymbol{\Theta}}(k)=\left(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}\right)\boldsymbol{\Theta}(k-1)$, 令 $\boldsymbol{\Theta}(k)=\dfrac{\tilde{\boldsymbol{\Theta}}(k)}{\sqrt{\tilde{\boldsymbol{\Theta}}(k)^{\mathrm{T}}\tilde{\boldsymbol{\Theta}}(k)}}$.

(3) 当 $\|\boldsymbol{\Theta}(k)-\boldsymbol{\Theta}(k-1)\|_\infty<\varepsilon$ 时, 则 $\boldsymbol{\Theta}^*=\boldsymbol{\Theta}(k)$, 否则令 $k=k+1$, 转 (2).

$\boldsymbol{\Theta}^*$ 求出以后, 把它代入式 (9.1.7) 即得最优组合赋权系数向量 $\boldsymbol{W}_c^*=\boldsymbol{W}\boldsymbol{\Theta}^*$.

由于传统的加权向量一般都是满足归一化约束条件, 因此, 为了与人们的习惯用法保持一致, 还需要对 $\boldsymbol{W}_c^*=(w_{c1}^*,w_{c2}^*,\cdots,w_{cn}^*)^{\mathrm{T}}$ 进行归一化处理. 即令

$$w_{cj}^{**}=w_{cj}^*\Big/\sum_{j=1}^n w_{cj}^*,\quad j=1,2,\cdots,n. \tag{9.1.22}$$

当然也可以先对 $\theta_k^*, k=1,2,\cdots,l$ 进行归一化处理, 即令

$$\theta_k^{**}=\theta_k^*\Big/\sum_{k=1}^l\theta_k^*,\quad k=1,2,\cdots,l. \tag{9.1.23}$$

再按 $\boldsymbol{W}_c^{**}=\boldsymbol{W}\boldsymbol{\Theta}^{**}$ 计算加权向量.

容易证明这两种归一化方法等价的. 于是, 得到的 $\boldsymbol{W}_c^*=(w_{c1}^*,w_{c2}^*,\cdots,w_{cn}^*)^{\mathrm{T}}$ 就是所求的满足归一化约束条件最优组合赋权系数向量. 实际上, 对 $\boldsymbol{W}_c^*$ 进行归一化处理并不影响多属性决策问题排序的结果. 这可从如下定理 9.1.2 看出.

定理 9.1.2[115] 分别使用组合赋权系数向量 $\boldsymbol{W}_c^*$ 和 $\boldsymbol{W}_c^{**}$ 按式 (9.1.9) 对多属性决策问题进行决策方案的排序, 则所得到的两个排序结果是相同的, 即若 $D_i(\boldsymbol{W}_c^*)>D_j(\boldsymbol{W}_c^*)$, 则有

$$D_i(\boldsymbol{W}_c^{**})>D_j(\boldsymbol{W}_c^{**}),\quad i,j=1,2,\cdots,m.$$

证明 由式 (9.1.9) 知

$$D_i(\boldsymbol{W}_c^*)=\sum_{j=1}^n b_{ij}w_{cj}^*=\sum_{j=1}^n b_{ij}\sum_{k=1}^l\theta_k^*w_{jk}=\sum_{k=1}^l\theta_k^*\sum_{j=1}^n b_{ij}w_{jk}.$$

同理 $D_i(\boldsymbol{W}_c^{**}) = \sum\limits_{k=1}^{l} \theta_k^{**} \sum\limits_{j=1}^{n} b_{ij} w_{jk}$, 由式 (9.1.23) 知

$$\theta_k^{**} = \theta_k^* \Big/ \sum_{k=1}^{l} \theta_k^*, \quad k = 1, 2, \cdots, l.$$

所以 $D_i(\boldsymbol{W}_c^{**}) = \sum\limits_{k=1}^{l} \theta_k^* \sum\limits_{j=1}^{n} b_{ij} w_{jk} \Big/ \sum\limits_{k=1}^{l} \theta_k^*$, 即

$$D_i(\boldsymbol{W}_c^{**}) = D_i(\boldsymbol{W}_c^*) \Big/ \sum_{k=1}^{l} \theta_k^*, \quad i = 1, 2, \cdots, m.$$

因为 $\sum\limits_{k=1}^{l} \theta_k^* > 0$, 则当 $D_i(\boldsymbol{W}_c^*) > D_j(\boldsymbol{W}_c^*)$ 时, 有

$$D_i(\boldsymbol{W}_c^{**}) > D_j(\boldsymbol{W}_c^{**}), \quad i, j = 1, 2, \cdots, m. \text{ 证毕.}$$

综上所述, 最优组合赋权系数向量的计算步骤可归纳为

(1) 由属性矩阵 $\boldsymbol{A} = (a_{ij})_{m\times n}$, 根据式 (9.1.1)~ (9.1.4) 计算规范化的属性矩阵 $\boldsymbol{B} = (b_{ij})_{m\times n}$, 以及根据定义式计算 n 阶对称非负定方阵 $\boldsymbol{B}_1$.

(2) 由 l 种具体不同的赋权方法组成的权系数向量矩阵 $\boldsymbol{W}$ 和 $\boldsymbol{B}_1$ 计算对称矩阵 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$, 同时计算 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 的最大特征根 $\lambda_{\max}$ 及其所对应的单位化特征向量 $\boldsymbol{\Theta}^*$.

(3) 根据 $\boldsymbol{W}_c^* = \boldsymbol{W}\boldsymbol{\Theta}^*$ 并进行归一化处理, 求出后由式 (9.1.22) 或式 (9.1.23) 得归一化的最优组合赋权向量 $\boldsymbol{W}_c^{**}$.

(4) 把 $\boldsymbol{W}_c^{**}$ 代入式 (9.1.9) 计算第 i 个决策方案 S_i 的多属性综合评价值 $D_i(\boldsymbol{W}_c^{**})$, i =1,2,$\cdots$,m, 并根据各决策方案多属性综合评价值 $D_i(\boldsymbol{W}_c^{**})$ 的大小对多属性决策问题进行排序, 作出科学的分析.

9.1.3 基于离差平方和的最优组合赋权方法的实例分析

下面以文献 [112] 中的例子的数据, 给出了本节提出的基于离差平方和的最优组合赋权方法的多属性决策问题的分析结果.

考虑一个购买住房的多属性决策问题. 其方案集 S ={S_1, S_2, S_3, S_4}, 分别表示 4 处可供选择的住房, 其属性集 P ={P_1, P_2, P_3, P_4, P_5}, 它们分别是购房价格 (万元)、使用面积 (平方米)、住房距工作地点的距离 (公里)、住户设施 (分数) 和住房周围环境 (分数) , 其中 P_1, P_3 为成本型属性,P_2, P_4, P_5 为效益型属性. 该问题

的原始决策矩阵 $\boldsymbol{A}$ 为

$$\boldsymbol{A}=\begin{bmatrix}30 & 100 & 10 & 7 & 7\\ 25 & 80 & 8 & 3 & 5\\ 18 & 50 & 20 & 5 & 10\\ 22 & 70 & 12 & 5 & 9\end{bmatrix}.$$

根据式 (9.1.1)、(9.1.2) 计算规范化的属性矩阵 $\boldsymbol{B}$ 为

$$\boldsymbol{B}=\begin{bmatrix}0 & 1 & 0.8333 & 1 & 0.4\\ 0.4167 & 0.6 & 1 & 0 & 0\\ 1 & 0 & 0 & 0.5 & 1\\ 0.6667 & 0.4 & 0.6667 & 0.5 & 0.8\end{bmatrix}.$$

根据定义式计算 5 阶对称非负定方阵 $\boldsymbol{B}_1$ 为

$$\boldsymbol{B}_1=\begin{bmatrix}4.2639 & -4.2000 & -3.5275 & -1.6668 & 3.0999\\ -4.2000 & 4.1600 & 3.5998 & 1.6000 & -3.0400\\ -3.5275 & 3.5998 & 4.6110 & -0.6668 & -4.0666\\ -1.6668 & 1.6000 & -0.6668 & 4.0000 & 1.6000\\ 3.0999 & -3.0400 & -4.0666 & 1.6000 & 4.7200\end{bmatrix}.$$

假设第一种赋权方法为主观赋权方法, 决策者事先给出的主观赋权向量为

$$\boldsymbol{W}_1=(0.3,0.2,0.15,0.15,0.2)^{\mathrm{T}}.$$

假设第二种赋权方法为客观赋权方法, 运用文献 [123] 给出的客观赋权法, 可以求出属性的权重向量为

$$\boldsymbol{W}_2=(0.1911,0.1824,0.2435,0.1849,0.1981)^{\mathrm{T}}.$$

即两个权系数向量组成的分块矩阵 $\boldsymbol{W}=(\boldsymbol{W}_1,\boldsymbol{W}_2),\boldsymbol{W}$ 实际上为 5×2 的矩阵.

由 $\boldsymbol{W}$ 和 $\boldsymbol{B}_1$ 计算对称矩阵 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 为

$$\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}=\begin{bmatrix}0.2240 & 0.1663\\ 0.1663 & 0.2582\end{bmatrix}.$$

从而计算出 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{W}$ 的最大特征根 $\lambda_{\max}=0.4082$, 它所对应的单位化特征向量

$$\boldsymbol{\Theta}^*=[0.6700,0.7424]^{\mathrm{T}}.$$

根据 $\boldsymbol{W}_c^* = \mathbf{W}\boldsymbol{\Theta}^*$ 算出

$$\boldsymbol{W}_c^* = 0.6700\boldsymbol{W}_1 + 0.7424\boldsymbol{W}_2 = (0.3429 \quad 0.2694 \quad 0.2813 \quad 0.2378 \quad 0.2811)^{\mathrm{T}}.$$

将 $\boldsymbol{W}_c^*$ 进行归一化处理得

$$\boldsymbol{W}_c^{**} = (0.2428 \quad 0.1907 \quad 0.1991 \quad 0.1683 \quad 0.1990)^{\mathrm{T}}.$$

$\boldsymbol{W}_c^*$ 代入式 (9.1.9) 分别计算 4 个决策方案 S_1, S_2, S_3, S_4 的多属性综合评价值为

$$\begin{aligned} &D_1(\boldsymbol{W}_c^{**}) = 0.6045, \quad D_2(\boldsymbol{W}_c^{**}) = 0.4147, \\ &D_3(\boldsymbol{W}_c^{**}) = 0.5259, \quad D_4(\boldsymbol{W}_c^{**}) = 0.6142. \end{aligned}$$

计算结果表明 $S_4 \succ S_1 \succ S_3 \succ S_2$, 其中 “$\succ$” 表示优于.

对此购买住房的多属性决策问题, 本节提出的方法所得的排序结果和文献 [112] 是一致的. 实例就说明了本节提出的以离差平方和为准则的最优组合赋权方法的有效性. 它综合主客观综合赋权法的特点, 同时相应的决策排序结果反映主观程度和客观程度.

9.1.4 基于离差最大化准则下的多属性决策的最优组合赋权方法 [116]

前已述及, 根据简单线性加权法, 由组合赋权系数向量 $\boldsymbol{W}_c$ 计算而得的第 i 个决策方案 S_i 的多属性综合评价值 $D_i(\boldsymbol{W}_c)$ 总是愈大愈好, $D_i(\boldsymbol{W}_c)$ 愈大表示第 i 个决策方案 S_i 愈优. 当组合赋权系数向量 $\boldsymbol{W}_c$ 已知的时候, 可根据 (9.1.9) 式计算出各个决策方案的多属性综合评价值, 从而可以对各决策方案进行排序. 然而组合赋权系数向量一般是未知的. 下面进一步讨论 $\boldsymbol{W}_c$ 的另外一种确定方法. 众所周知, 在多属性决策中, 如果第 j 个属性 P_j 对所有决策方案而言均无差别, 则属性 P_j 对决策方案的排序将不起作用, 这样的属性可令其权系数为 0; 反之, 如果 P_j 使所有决策方案的属性值有较大差异, 这样的属性对决策方案的排序将起较大作用, 此时应该给 P_j 赋予较大的权系数. 在统计学中, 离差是反映差异程度的一个重要指标. 基于这样的原理, 选择组合赋权系数向量的另一个基本思想是使所有 n 个属性对所有 m 个决策方案的总离差达到最大 [122]. 为此, 定义第 i 个决策方案和其他决策方案关于属性 P_j 的离差之和 $v_{ij}(\boldsymbol{W}_c)$ 为

$$\begin{aligned} v_{ij}(\boldsymbol{W}_c) = \sum_{i_1=1}^{m} |b_{ij}w_{cj} - b_{i_1 j}w_{cj}| = \sum_{i_1=1}^{m} |b_{ij} - b_{i_1 j}|\, w_{cj}, \\ i = 1, 2, \cdots, m, \quad j = 1, 2, \cdots, n. \end{aligned} \tag{9.1.24}$$

对于属性 P_j 来说, 设 $v_j(\boldsymbol{W}_c)$ 表示所有 m 个决策方案与其他决策方案之总离差, 则有

$$v_j(\boldsymbol{W}_c)=\sum_{i=1}^{m} v_{ij}(\boldsymbol{W}_c)=\sum_{i=1}^{m}\sum_{i_1=1}^{m}|b_{ij}-b_{i_1j}|\,w_{cj},\quad j=1,2,\cdots,n. \tag{9.1.25}$$

根据前述的选择组合赋权系数向量 $\boldsymbol{W}_c$ 的基本思想, 应该使所有 n 个属性对所有 m 个决策方案的总离差达到最大, 这样有利于决策方案的排序. 于是可构造如下目标函数

$$J_2(\boldsymbol{W}_c)=\sum_{j=1}^{n} v_j(\boldsymbol{W}_c)=\sum_{j=1}^{n}\sum_{i=1}^{m}\sum_{i_1=1}^{m}|b_{ij}-b_{i_1j}|\,w_{cj}. \tag{9.1.26}$$

若令 n 维行向量 $\boldsymbol{B}_1$ 为

$$\boldsymbol{B}_1=\left[\sum_{i=1}^{m}\sum_{i_1=1}^{m}|b_{i1}-b_{i_11}|\ \ \sum_{i=1}^{m}\sum_{i_1=1}^{m}|b_{i2}-b_{i_12}|\ \cdots\ \sum_{i=1}^{m}\sum_{i_1=1}^{m}|b_{in}-b_{i_1n}|\right]. \tag{9.1.27}$$

则目标函数 $J_2(\boldsymbol{W}_c)$ 可表为

$$J_2(\boldsymbol{W}_c)=\boldsymbol{B}_1\boldsymbol{W}_c, \tag{9.1.28}$$

其中 $\boldsymbol{W}_c=(w_{c1},w_{c2},\cdots,w_{cn})^{\mathrm{T}}$ 为组合赋权系数向量.

类似地, 要想求出组合赋权系数向量 $\boldsymbol{W}_c$, 由式 (9.1.9) 知, 只需求出组合权系数的线性表出系数向量 $\boldsymbol{\Theta}$. 既然 $\boldsymbol{W}_c$ 为 $\boldsymbol{\Theta}$ 的函数. 从而式 (9.1.28) 表明目标函数 $J_2(\boldsymbol{W}_c)$ 也为 $\boldsymbol{\Theta}$ 的函数, 此时 $J_2(\boldsymbol{W}_c)$ 可记为 $G(\boldsymbol{\Theta})$. 于是, 基于离差最大化的多属性决策的组合赋权方法表为如下最优化问题

$$\begin{aligned}&\max G(\boldsymbol{\Theta})=\boldsymbol{B}_1\mathbf{W}\boldsymbol{\Theta},\\&\text{s.t.}\begin{cases}\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta}=1,\\ \boldsymbol{\Theta}\geqslant \mathbf{0},\end{cases}\end{aligned} \tag{9.1.29}$$

其中 $\boldsymbol{W}=(\boldsymbol{W}_1,\boldsymbol{W}_2,\cdots,\boldsymbol{W}_l)$ 为 l 种具体的赋权方法组成的分块矩阵.

对于最优化问题 (9.1.29) 的解的讨论, 具有如下结论.

定理 9.1.3[116]　模型 (9.1.29) 所确定的最优化模型的最优解为

$$\theta_k^*=\boldsymbol{B}_1\boldsymbol{W}_k\Bigg/\sqrt{\sum_{k=1}^{l}(\boldsymbol{B}_1\boldsymbol{W}_k)^2},\quad k=1,2,\cdots,l. \tag{9.1.30}$$

证明　构造 Lagrange 函数

$$L(\theta_1,\theta_2,\cdots,\theta_l)=\boldsymbol{B}_1\boldsymbol{W}\boldsymbol{\Theta}+\lambda(\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta}-1)=\boldsymbol{B}_1\sum_{k=1}^{l}\boldsymbol{W}_k\theta_k+\lambda\left(\sum_{k=1}^{l}\theta_k^2-1\right),$$

其中 λ 为 Lagrange 乘子, 令 $\dfrac{\partial L}{\partial \theta_k}=0$ 得

$$\boldsymbol{B}_1\boldsymbol{W}_k+2\lambda\theta_k=0,\quad k=1,2,\cdots,l.$$

所以

$$\theta_k=-\boldsymbol{B}_1\boldsymbol{W}_k/2\lambda,\quad k=1,2,\cdots,l. \tag{9.1.31}$$

把式 (9.1.31) 代入模型 (9.1.2) 的约束条件中得

$$\lambda=-\sqrt{\sum_{k=1}^{l}(\boldsymbol{B}_1\boldsymbol{W}_k)^2}\bigg/2.$$

所以 $\theta_k^*=\boldsymbol{B}_1\boldsymbol{W}_k\bigg/\sqrt{\sum\limits_{k=1}^{l}(\boldsymbol{B}_1\boldsymbol{W}_k)^2},k=1,2,\cdots,l$ 证毕.

将式 (9.1.30) 代入式 (9.1.7) 得到基于离差最大化的多属性决策的最优组合赋权向量为

$$\boldsymbol{W}_c^*=\boldsymbol{W}\boldsymbol{\Theta}^*=\sum_{k=1}^{l}\theta_k\boldsymbol{W}_k. \tag{9.1.32}$$

对 $\boldsymbol{W}_c^*=(w_{c1}^*,w_{c2}^*,\cdots,w_{cn}^*)^{\mathrm{T}}$ 还需要进行归一化处理. 此时可按式 (9.1.17) 对 $\theta_k^*,k=1,2,\cdots,l$ 进行归一化处理即可. 将式 (9.1.17) 代入式 (9.1.7) 得基于离差最大化的多属性决策的最优非负归一化组合赋权向量为

$$\boldsymbol{W}_c^{**}=\sum_{k=1}^{l}\theta_k^{**}\boldsymbol{W}_k. \tag{9.1.33}$$

由定理 9.1.2 知, 实际上对组合赋权系数向量进行归一化处理并不影响决策方案的排序结果.

基于离差最大化准则下的多属性决策的最优组合赋权系数向量的计算步骤为

(1) 由属性矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$, 根据式 (9.1.1) ~ (9.1.4) 计算规范化的属性矩阵 $\boldsymbol{B}=(b_{ij})_{m\times n}$, 以及根据式 (9.1.27) 计算 n 维行向量 $\boldsymbol{B}_1$.

(2) 计算 n 维行向量 $\boldsymbol{B}_1$ 和第 k 种赋权方法给出的权系数向量 $\boldsymbol{W}_k$ 的乘积 $\boldsymbol{B}_1\boldsymbol{W}_k,k=1,2,\cdots,l$, 根据式 (9.1.30) 计算最优组合赋权向量的线性表出系数 $\boldsymbol{\Theta}^*=(\theta_1^*,\theta_2^*,\cdots,\theta_l^*)^{\mathrm{T}}$.

(3) 根据式 (9.1.17) 计算 $\boldsymbol{\Theta}^*$ 的归一化向量 $\boldsymbol{\Theta}^{**}=(\theta_1^{**},\theta_2^{**},\cdots,\theta_l^{**})^{\mathrm{T}}$, 再根据式 (9.1.33) 求出归一化的最优组合赋权向量 $\boldsymbol{W}_c^{**}$.

(4) 把 $\boldsymbol{W}_c^{**}$ 代入式 (9.1.9) 计算第 i 个决策方案 S_i 的多属性综合评价值 $D_i(\boldsymbol{W}_c^{**})$, i =1,2,$\cdots,m$, 并根据各决策方案多属性综合评价值 $D_i(\boldsymbol{W}_c^{**})$ 的大小对多属性决策问题进行排序.

9.1.5 基于离差最大化准则下的最优组合赋权方法的实例分析

下面以文献 [113] 中的例子的数据, 给出了本节提出的基于离差最大化的多属性决策的最优组合赋权向量的分析结果.

这是一个在市场上选择机器人的多属性决策问题. 考虑一个用户要选择机器人, 其方案集为 $S=\{S_1,S_2,S_3,S_4\}$, 分别表示 4 个可供选择的方案; 其属性集 $P=\{P_1,P_2,P_3,P_4\}$, 即有 4 个属性, 它们分别是, P_1 表示价格 ($ 10000); P_2 表示速度 (m/s) ; P_3 表示可重复性 (mm) ; P_4 表示负载能力 (kg). 其中 P_1, P_3 为成本型属性, P_2, P_4 为效益型属性. 该问题的原始决策矩阵为

$$\boldsymbol{A}=\begin{bmatrix} 3.0 & 1.0 & 1.0 & 70 \\ 2.5 & 0.8 & 0.8 & 50 \\ 1.8 & 0.5 & 2.0 & 110 \\ 2.2 & 0.7 & 1.2 & 90 \end{bmatrix}.$$

根据式 (9.1.1)、(9.1.2) 计算规范化的属性矩阵为

$$\boldsymbol{B}=\begin{bmatrix} 0 & 1 & 0.8333 & 0.3333 \\ 0.4167 & 0.6 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0.6667 & 0.4 & 0.6667 & 0.6667 \end{bmatrix}.$$

根据定义式计算 4 维行向量为

$$\boldsymbol{B}_1=[6.5000 \quad 6.4000 \quad 6.3332 \quad 6.6666].$$

假设第一类赋权方法为主观赋权方法, 机器人用户聘请 3 个专家给出关于属性的权重向量分别为

$$\begin{aligned} \boldsymbol{W}_1 &= (0.3, 0.4, 0.15, 0.15)^{\mathrm{T}}, \\ \boldsymbol{W}_2 &= (0.4, 0.3, 0.15, 0.15)^{\mathrm{T}}, \\ \boldsymbol{W}_3 &= (0.25, 0.25, 0.25, 0.25)^{\mathrm{T}}, \end{aligned}$$

假设第二类赋权方法为客观赋权方法, 运用文献 [123] 给出的客观赋权法, 可以求出 4 个属性的权重向量为

$$\boldsymbol{W}_2=(0.2403 \quad 0.2294 \quad 0.3062 \quad 0.2242)^{\mathrm{T}}.$$

计算 4 维行向量 $\boldsymbol{B}_1$ 和第 k 种赋权方法给出的权系数向量 $\boldsymbol{W}_k$ 的乘积, $k=1,2,3,4$.

$$\boldsymbol{B}_1\boldsymbol{W}_1=7.6953, \quad \boldsymbol{B}_1\boldsymbol{W}_2=7.6326, \quad \boldsymbol{B}_1\boldsymbol{W}_3=5.5596, \quad \boldsymbol{B}_1\boldsymbol{W}_4=5.0130.$$

根据式 (9.1.17) 计算

$$\boldsymbol{\Theta}^{**}=(\theta_1^{**},\theta_2^{**},\theta_3^{**},\theta_4^{**})^{\mathrm{T}}=(0.2971\quad 0.2947\quad 0.2147\quad 0.1935)^{\mathrm{T}}.$$

再根据式 (9.1.33) 求出归一化的最优组合赋权向量

$$\begin{aligned}\boldsymbol{W}_c^{**}=\sum_{k=1}^{4}\theta_k^{**}\boldsymbol{W}_k&=0.2971\boldsymbol{W}_1+0.2947\boldsymbol{W}_2+0.2147\boldsymbol{W}_3+0.1935\mathrm{W}_4\\&=(0.3072\quad 0.3053\quad 0.2017\quad 0.1858)^{\mathrm{T}}.\end{aligned}$$

把 $\boldsymbol{W}_c^{**}$ 代入式 (9.1.9) 分别计算第 i 个决策方案 S_i 的多属性综合评价值

$$D_1(\boldsymbol{W}_c^{**})=0.5353,\quad D_2(\boldsymbol{W}_c^{**})=0.5129,$$
$$D_3(\boldsymbol{W}_c^{**})=0.4930,\quad D_4(\boldsymbol{W}_c^{**})=0.6142.$$

计算结果表明 $S_4\succ S_1\succ S_2\succ S_3$ ，其中 "$\succ$" 表示优于.

对市场上选择机器人的多属性决策问题, 本节提出的方法所得的排序结果和文献 [113] 是一致的. 实例说明以离差平方和为准则的最优组合赋权方法的有效性. 它综合了各种赋权方法的特点, 通过一个最优化数学模型求出组合赋权系数. 同时, 相应的决策排序结果反映主观程度和客观程度, 具有概念清楚、涵义明确的特点, 同时计算也不复杂, 应用举例表明本方法排序结果准确性.

9.2 最优证券组合投资决策动态模型研究

随着人们收入水平的增加, 投资方式正在向多元化方向发展. 证券投资常常是投资者愿意选择的主要投资方式之一. 在有价证券市场上, 投资者最关心的是投资收益的大小以及投资风险的高低. 一般来说, 证券投资收益往往与投资风险成正比关系, 即投资收益越大, 投资风险也越大. 由诺贝尔奖得主 Markowitz 首先提出的证券组合投资模型奠定了现代投资理论的基础 [118]. 然而, 目前大多数证券组合投资决策模型存在的问题表现在两个方面：一是得出的组合投资的权向量是个常向量, 它不随时间变化而变化. 实际上不同时点上证券组合投资的权向量不可能相等. 二是证券组合投资的权向量是根据已经发生的历史资料由模型计算得出的, 用它作为未来时刻的证券组合投资决策的权向量实为 "马后炮" , 没有实际意义. 正是基于上面的两点考虑, 本节提出了在两种不同风险度量指标下的动态证券组合投资决策模型, 并给出相应的规划论算法.

9.2.1 以方差作为风险度量指标的证券组合投资动态模型 [117]

设证券市场上有 n 种证券, 则第 i 种证券第 t 期单位投资额的收益率 β_{it} 为

$$\beta_{it}=(p_{it}^1-p_{it}^0+d_{it})/p_{it}^0,\quad i=1,2,\cdots,n,\quad t=1,2,\cdots,N,\tag{9.2.1}$$

其中 p_{it}^1 表示第 i 种证券第 t 期出售价格, p_{it}^0 表示相应的买入价, d_{it} 表示相应的持有期所获得的红利、股息等. 显然 β_{it} 为一随机变量.

假设投资者在第 t 期投资金额为 m_t, m_t 为外生变量, 令 x_{it} 表示投资者投资到第 i 种证券第 t 期投资额, x_{it} 为内生变量.

我们采用证券收益率的数学期望作为证券收益大小的度量指标, 用证券收益率的方差作为风险度量指标, 记

$$b_{it} = E\beta_{it}, \quad \sigma_{ij}(t) = E(\beta_{it} - b_{it})(\beta_{jt} - b_{jt}), \tag{9.2.2}$$

其中 b_{it} 表示 β_{it} 的数学期望, $\sigma_{ij}(t)$ 表示第 i 种证券与第 j 种证券收益率的协方差.

在上述记号条件下, 第 t 期证券组合投资总的期望收益 b_t 为

$$b_t = \sum_{i=1}^{n} b_{it}x_{it}, \quad t = 1, 2, \cdots, N. \tag{9.2.3}$$

第 t 期 n 种证券组合投资总的风险 σ_t^2 为

$$\sigma_t^2 = E\left(\sum_{i=1}^{n} \beta_{it}x_{it} - \sum_{i=1}^{n} b_{it}x_{it}\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij}(t)x_{it}x_{jt}, \quad t = 1, 2, \cdots, N. \tag{9.2.4}$$

作为理性的投资者, 一般是相对风险规避型的. 应选择在一定收益的条件下使得投资风险最小, 即满足以下几个条件:

(1) 设 r_t 为投资者心目中最低的收益率, 如 r_t 可视为同期银行利率, 则组合证券投资的预期收益应大于或等于在银行储蓄存款所获得的收益. 即

$$\sum_{i=1}^{n} b_{it}x_{it} \geqslant r_t m_t, \quad t = 1, 2, \cdots, N. \tag{9.2.5}$$

(2) 第 t 期投资到 n 种证券的资金总量受到投资者拥有的资金的约束. 即

$$\sum_{i=1}^{n} x_{it} = m_t, \quad t = 1, 2, \cdots, N. \tag{9.2.6}$$

(3) 随着证券法的健全, 现实证券市场上卖空操作不被允许. 即

$$x_{it} \geqslant 0, \quad i = 1, 2, \cdots, n, \quad t = 1, 2, \cdots, N. \tag{9.2.7}$$

(4) 投资者是风险规避的, 要求 N 期时间内组合证券投资方差达到最小, 则有如下目标函数

$$\min \sigma^2 = \sum_{t=1}^{N}\sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij}(t)x_{it}x_{jt}.$$

综合上面四个条件, 可提出如下证券组合投资的动态模型

$$\min \sigma^2 = \sum_{t=1}^{N}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}(t)x_{it}x_{jt},$$
$$\text{s.t.}\begin{cases}\sum_{i=1}^{n}x_{it}=m_t, t=1,2,\cdots,N,\\ \sum_{i=1}^{n}b_{it}x_{it}\geqslant r_t m_t, t=1,2,\cdots,N,\\ x_{it}\geqslant 0, i=1,2,\cdots,n, t=1,2,\cdots,N.\end{cases} \tag{9.2.8}$$

若令 $\boldsymbol{X}_t=(x_{1t},x_{2t},\cdots,x_{nt})^{\mathrm{T}}$, $\boldsymbol{X}=(\boldsymbol{X}_1^{\mathrm{T}},\boldsymbol{X}_2^{\mathrm{T}},\cdots,\boldsymbol{X}_N^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{V}_t=(\sigma_{ij}(t))_{n\times n}$, 一般假定 $\boldsymbol{V}_t$ 为 n 阶正定阵. $\boldsymbol{I}$ 为元素 1 构成的 n 维列向量, 即 $\boldsymbol{I}=(1,1,\cdots,1)^{\mathrm{T}}$, $\boldsymbol{B}_0=(r_1m_1,r_2m_2,\cdots,r_Nm_N)^{\mathrm{T}}$, $\boldsymbol{B}_t=(b_{1t},b_{2t},\cdots,b_{nt})^{\mathrm{T}}$, $t=1,2,\cdots,N$, $\boldsymbol{M}=(m_1,m_2,\cdots,m_N)^{\mathrm{T}}$. 再令

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{I}^{\mathrm{T}} & & & \\ & \boldsymbol{I}^{\mathrm{T}} & & \\ & & \ddots & \\ & & & \boldsymbol{I}^{\mathrm{T}}\end{bmatrix}_{N\times Nn},\quad \boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1^{\mathrm{T}} & & & \\ & \boldsymbol{B}_2^{\mathrm{T}} & & \\ & & \ddots & \\ & & & \boldsymbol{B}_N^{\mathrm{T}}\end{bmatrix}_{N\times Nn},$$

$$\boldsymbol{V}=\begin{bmatrix}\boldsymbol{V}_1 & & & \\ & \boldsymbol{V}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{V}_n\end{bmatrix}_{Nn\times Nn},$$

其中 $\boldsymbol{A},\boldsymbol{B}$ 分别为 $N\times Nn$ 的分块对角阵, $\boldsymbol{V}$为 $Nn\times Nn$ 的分块对角阵. $\boldsymbol{X}$ 为 Nn 维列向量, 即为证券组合投资的权向量. 则上述模型 (9.2.1) 可表为如下矩阵形式

$$\min \sigma^2=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{X},$$
$$\text{s.t.}\begin{cases}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{M},\\ \boldsymbol{B}\boldsymbol{X}\geqslant\boldsymbol{B}_0,\\ \boldsymbol{X}\geqslant\boldsymbol{0}.\end{cases} \tag{9.2.9}$$

模型 (9.2.9) 是一个凸二次规划问题, Kuhn-Tucker 条件是该问题最优点存在

的充要条件, 其 Kuhn-Tucker 条件可表为

$$\begin{cases} 2\boldsymbol{V}\boldsymbol{X}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Lambda}-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{U}-\boldsymbol{Y}=\boldsymbol{0}, \\ \boldsymbol{A}\boldsymbol{X}=\boldsymbol{M}, \\ \boldsymbol{B}\boldsymbol{X}-\boldsymbol{B}_0-\boldsymbol{Z}=\boldsymbol{0}, \\ \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{0},\boldsymbol{U}^{\mathrm{T}}\boldsymbol{Z}=\boldsymbol{0}, \\ \boldsymbol{X}\geqslant\boldsymbol{0},\boldsymbol{Y}\geqslant\boldsymbol{0},\boldsymbol{Z}\geqslant\boldsymbol{0},\boldsymbol{U}\geqslant\boldsymbol{0}, \end{cases} \tag{9.2.10}$$

其中 $\boldsymbol{\Lambda}=(\lambda_1,\lambda_2,\cdots,\lambda_N)^{\mathrm{T}}$ 是与模型 (9.2.9) 约束条件中第一式相对应的 Langrange 乘子向量, $\boldsymbol{Y}$ 是与模型 (9.2.9) 约束条件中第二式相对应的 Kuhn-Tucker 乘子向量, $\boldsymbol{Y}$ 为 Nn 维列向量, $\boldsymbol{U}$ 是与模型 (9.2.9) 约束条件中第三式相对应的 Kuhn-Tucker 乘子向量, $\boldsymbol{U}$ 为 N 维列向量, $\boldsymbol{Z}$ 是与模型 (9.2.9) 约束条件中第二式相对应的松弛变量组成的 N 维列向量. 为求满足上述 Kuhn-Tucker 条件的解, 可考虑如下线性规划问题

$$\min J=\boldsymbol{e}_N^{\mathrm{T}}\boldsymbol{W},$$
$$\text{s.t.}\begin{cases} 2\boldsymbol{V}\boldsymbol{X}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Lambda}-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{U}-\boldsymbol{Y}=\boldsymbol{0}, \\ \boldsymbol{A}\boldsymbol{X}+\boldsymbol{W}=\boldsymbol{M}, \\ \boldsymbol{B}\boldsymbol{X}-\boldsymbol{B}_0-\boldsymbol{Z}=\boldsymbol{0}, \\ \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{0},\boldsymbol{U}^{\mathrm{T}}\boldsymbol{Z}=\boldsymbol{0}, \\ \boldsymbol{X}\geqslant\boldsymbol{0},\boldsymbol{Y}\geqslant\boldsymbol{0},\boldsymbol{Z}\geqslant\boldsymbol{0},\boldsymbol{U}\geqslant\boldsymbol{0},\boldsymbol{W}\geqslant\boldsymbol{0}, \end{cases} \tag{9.2.11}$$

其中 $\boldsymbol{e}_N^{\mathrm{T}}=(1,1,\cdots,1)_{1\times N}^{\mathrm{T}}$, $\boldsymbol{W}$ 为人工变量组成的 N 维列向量, 若能求出模型 (9.2.11) 的最优解, 且最优值 J 为 0, 则模型 (9.2.11) 最优解中部分向量 $\boldsymbol{X}^*$ 即为模型 (9.2.9) 的最优解.

9.2.2 以绝对离差作为风险度量指标的组合证券投资动态模型 [117]

若以方差作为风险度量指标, 我们必须计算 N 期协方差矩阵 $\boldsymbol{V}$, 当 n 和 N 较大时, 实际计算 $\boldsymbol{V}$ 时较为繁琐, 为此可考虑以绝对离差作为风险度量指标. 设 $J(\boldsymbol{X})$ 表示 N 期组合证券投资的绝对离差, 则 $J(\boldsymbol{X})$ 可表为

$$J(\boldsymbol{X})=\sum_{t=1}^{N}\left|\sum_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}\right|, \tag{9.2.12}$$

其中 β_{it} 为第 i 种证券第 t 期单位投资额收益的观察值, $\overline{\beta_i}=\dfrac{1}{N}\displaystyle\sum_{t=1}^{N}\beta_{it}$.

类似地, 可建立以绝对离差作为风险的组合证券投资动态模型

$$\min J(\boldsymbol{X})=\sum_{t=1}^{N}\left|\sum_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}\right|,$$

$$\text{s.t.}\begin{cases}\sum\limits_{i=1}^{n}x_{it}=m_t,t=1,2,\cdots,N,\\ \sum\limits_{i=1}^{n}\beta_{it}x_{it}\geqslant r_tm_t,t=1,2,\cdots,N,\\ x_{it}\geqslant 0,i=1,2,\cdots,n,t=1,2,\cdots,N.\end{cases}\tag{9.2.13}$$

为求解模型 (9.2.13) , 令

$$w_1(t)=\begin{cases}\sum\limits_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}, & \text{当}\sum\limits_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}\geqslant 0.\\ 0, & \text{当}\sum\limits_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}\leqslant 0,\end{cases}\tag{9.2.14}$$

$$w_2(t)=\begin{cases}-\sum\limits_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}, & \text{当}\sum\limits_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}\leqslant 0,\\ 0, & \text{当}\sum\limits_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}\geqslant 0.\end{cases}\tag{9.2.15}$$

则

$$\left|\sum_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}\right|=w_1(t)+w_2(t),\quad \sum_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}=w_1(t)+w_2(t).$$

模型 (9.2.13) 可通过如下线性规划模型来求解

$$\min J(x)=\sum_{t=1}^{N}[w_1(t)+w_2(t)],$$

$$\text{s.t.}\begin{cases}\sum\limits_{i=1}^{n}(\beta_{it}-\overline{\beta_i})x_{it}-w_1(t)+w_2(t)=0,t=1,2,\cdots,N,\\ \sum\limits_{i=1}^{n}x_{it}=m_t,t=1,2,\cdots,N,\\ \sum\limits_{i=1}^{n}\beta_{it}x_{it}\geqslant r_tm_t,t=1,2,\cdots,N,\\ x_{it}\geqslant 0,w_1(t)\geqslant 0,w_2(t)\geqslant 0,i=1,2,\cdots,n,t=1,2,\cdots,N.\end{cases}\tag{9.2.16}$$

模型 (9.2.16) 的求解可以利用现成的线性规划软件.

9.3 证券组合投资的多目标区间数线性规划模型

许多国内外学者依据这些理论对证券组合投资进行了深入研究, 采用不同的方法得出一系列的研究成果. 然而, 目前绝大多数证券组合投资模型是针对投资收益率和风险均为精确数的情形进行的研究. 实际上证券市场不确定性的信息使得组合投资的收益是一个动态波动值[119]. 因此, 针对投资收益率和风险均为区间数的组合投资模型的研究更具有实际意义.

有关区间数线性规划问题是国内外学者研究的热点问题之一[120~122], 并且对其求解方法提出了许多行之有效的方法[123,124]. 文献 [119] 提出了一种区间数的组合投资模型, 在此基础上, 本节建立了证券组合投资的多目标区间数线性规划模型, 通过引入收益–风险偏好参数和优化水平参数, 讨论了多目标区间数线性规划模型的有效解, 最后进行了实例分析.

9.3.1 证券组合投资的多目标区间数线性规划模型的建立[126]

若投资者选择 n 种风险证券进行组合投资, 设区间数 $r_{it}=[r_{it}^-,r_{it}^+]$ 表示第 i 种证券 S_i 在持有期 t 内投资收益的范围. 证券组合投资的风险用风险损失率来度量, 风险损失率是指一个投资周期内资产在发生风险时可能的损失在总投资中所占的百分比 [119]. 设区间数 $q_{it}=[q_{it}^-,q_{it}^+]$ 表示第 i 种证券 S_i 在持有期 t 内风险损失率的范围, x_{it} 表示投资者在持有期 t 内第 i 种证券 S_i 的投资比例.

在证券组合投资中, 投资者最美好的愿望就是追求组合投资收益最大化和组合投资风险最小化. 于是有模型

$$\max\sum_{i=1}^{n} r_{it}x_{it},\qquad \text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{n} x_{it}=1, x_{it}\geqslant 0,\\ i=1,2,\cdots,n, t=1,2,\cdots,m\end{cases}\tag{9.3.1}$$

和模型

$$\min\sum_{i=1}^{n} q_{it}x_{it}.\qquad \text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{n} x_{it}=1, x_{it}\geqslant 0,\\ i=1,2,\cdots,n, t=1,2,\cdots,m.\end{cases}\tag{9.3.2}$$

模型 (9.3.1) 表明投资者只关心投资收益, 只要投资收益最大而不管风险如何. 该投资者属于极度风险偏好型. 模型 (9.3.2) 表明投资者只关心投资风险, 只要投

资风险最小而不管收益如何, 该投资者属于极度风险厌恶型. 这是两种理想状态但不切合实际. 因为模型 (9.3.1) 和模型 (9.3.2) 的约束条件相同, 所以把它们看成两目标的区间数线性规划模型

$$\max R(x_t) = \sum_{i=1}^{n} r_{it}x_{it},$$
$$\max Q(x_t) = -\sum_{i=1}^{n} q_{it}x_{it},$$
$$\text{s.t.} \begin{cases} \sum\limits_{i=1}^{n} x_{it} = 1, x_{it} \geqslant 0, \\ i = 1, 2, \cdots, n, t = 1, 2, \cdots, m. \end{cases} \tag{9.3.3}$$

9.3.2 证券组合投资的多目标区间数线性规划模型的求解[126]

常用区间数的加法, 减法和数乘运算规则[125] 如下:

设 $A = [a^-, a^+], B = [b^-, b^+]$, 则

$$A + B = [a^- + b^-, a^+ + b^+], \quad A - B = [a^- - b^+, a^+ - b^-],$$

$$kA = \begin{cases} [ka^-, ka^+], & k \geqslant 0, \\ [ka^+, ka^-], & k < 0. \end{cases}$$

通过引入参数 $\alpha(0 \leqslant \alpha \leqslant 1)$ 将模型 (9.3.3) 转化为单目标区间数线性规划模型

$$\max Z(x_t) = \alpha\sum_{i=1}^{n} r_{it}x_{it} - (1-\alpha)\sum_{i=1}^{n} q_{it}x_{it},$$
$$\text{s.t.} \begin{cases} \sum\limits_{i=1}^{n} x_{it} = 1, x_{it} \geqslant 0, \\ i = 1, 2, \cdots, n, t = 1, 2, \cdots, m, \end{cases} \tag{9.3.4}$$

这里, 当 $\alpha = 0$ 时, $\max Z(x_t) = -\sum\limits_{i=1}^{n} q_{it}x_{it}$, 即 $\min Z'(x_t) = \sum\limits_{i=1}^{n} q_{it}x_{it}$, 表明投资者只关心风险; 当 $\alpha = 1$ 时, $\max Z(x_t) = \sum\limits_{i=1}^{n} r_{it}x_{it}$, 表明投资者只关心收益. 于是, 我们称 α 为风险偏好参数, 它的大小反映投资者对投资收益和投资风险的一种权衡.

因此, 将模型 (9.3.4) 中的目标函数整理如下:

$$\begin{aligned}\max Z(x_t) &= \alpha\sum_{i=1}^{n} r_{it}x_{it} - (1-\alpha)\sum_{i=1}^{n} q_{it}x_{it}\\ &= \sum_{i=1}^{n}\alpha[r_{it}^-, r_{it}^+]x_{it} - \sum_{i=1}^{n}(1-\alpha)[q_{it}^-, q_{it}^+]x_{it}\\ &= \sum_{i=1}^{n}[\alpha r_{it}^- - (1-\alpha)q_{it}^+, ar_{it}^+ - (1-\alpha)q_{it}^-]x_{it}.\end{aligned}$$

令 $c_{it} = [c_{it}^-, c_{it}^+] = [\alpha r_{it}^- - (1-\alpha)q_{it}^+,\quad \alpha r_{it}^+ - (1-\alpha)q_{it}^-]$, 则有模型 (9.3.4) 化为

$$\max Z(x_t) = \sum_{i=1}^{n} c_{it}x_{it},$$

$$\text{s.t.}\begin{cases}\displaystyle\sum_{i=1}^{n} x_{it} = 1, x_{it} \geqslant 0,\\ i = 1,2,\cdots,n, t = 1,2,\cdots,m.\end{cases} \tag{9.3.5}$$

此模型为目标函数为区间数的线性规划模型. 通过引入目标函数优化水平参数 $\beta(0 \leqslant \beta \leqslant 1)$, β 反映了投资者对金融市场的客观情况的乐观估计程度, β 越靠近 1 体现投资者越乐观, 模型 (9.3.5) 可化为一般参数线性规划模型

$$\begin{aligned}\max Z(x_t) &= \sum_{i=1}^{n}[\beta c_{it}^+ + (1-\beta)c_{it}^-]x_{it}\\ &= \sum_{i=1}^{n}\{\beta[\alpha r_{it}^+ - (1-\alpha)q_{it}^-] + (1-\beta)[\alpha r_{it}^- - (1-\alpha)q_{it}^+]\}x_{it},\end{aligned}$$

$$\text{s.t.}\begin{cases}\sum\limits_{i=1}^{n} x_{it} = 1, x_{it} \geqslant 0,\\ i = 1,2,\cdots,n, t = 1,2,\cdots,m.\end{cases} \tag{9.3.6}$$

此模型利用一般参数线性规划方法求出最优解 $x_t(\alpha,\beta) = (x_{1t}(\alpha,\beta), x_{2t}(\alpha,\beta),\cdots, x_{nt}(\alpha,\beta))$, 它是证券组合投资的一种有效方案. 投资者可根据自己对投资风险的偏好程度以及金融市场的客观情况, 适当的估计参数 α 和 β, 可使投资者的投资收益较大而投资风险较小.

9.3.3 证券组合投资的多目标区间数线性规划模型的实例分析[126]

这里引用文献 [119] 的数据. 2000 年 1~6 月份齐鲁石化、东北高速、武钢股份和东风汽车 4 种证券的收益和风险损失率范围整理成表 9.3.1.

表 9.3.1 持有期内 4 种风险证券收益和风险损失率区间变化值

证券名称	齐鲁石化	东北高速	武钢股份	东风汽车
收益波动	[−0.0138, 0.1343]	[−0.0258, 0.2767]	[0.0339, 0.1136]	[−0.034, 0.0867]
风险损失率	[0.023, 0.034]	[0.015, 0.026]	[0.011, 0.018]	[0.013, 0.022]

依据表 9.3.1 数据建立证券组合投资的多目标区间数线性规划模型

$$
\begin{aligned}
\max R(x_t) =&[-0.0138, 0.1343]x_{1t} + [-0.0258, 0.2767]x_{2t} \\
&+ [0.0339, 0.1136]x_{3t} + [-0.034, 0.0867]x_{4t}, \\
\min Q(x_t) =&[0.023, 0.034]x_{1t} + [0.015, 0.026]x_{2t} \\
&+ [0.011, 0.018]x_{3t} + [0.013, 0.022]x_{4t}, \\
\text{s.t.}&\begin{cases} x_{1t} + x_{2t} + x_{3t} + x_{4t} = 1, \\ x_{it} \geqslant 0, i = 1, 2, 3, 4. \end{cases}
\end{aligned} \tag{9.3.7}
$$

通过引入风险偏好系数 α 和目标函数优化水平参数 β, 将 (9.3.7) 转化为 (9.3.8) 得

$$
\begin{aligned}
\max Z(x_t) =&\{\beta[0.1343\alpha - 0.023(1-\alpha)] + (1-\beta)[-0.0138\alpha - 0.034(1-\alpha)]\}x_{1t} \\
&+ \{\beta[0.2767\alpha - 0.015(1-\alpha)] + (1-\beta)[-0.0258\alpha - 0.026(1-\alpha)]\}x_{2t} \\
&+ \{\beta[0.1136\alpha - 0.011(1-\alpha)] + (1-\beta)[-0.039\alpha - 0.018(1-\alpha)]\}x_{3t} \\
&+ \{\beta[0.0867\alpha - 0.013(1-\alpha)] + (1-\beta)[-0.0347\alpha - 0.022(1-\alpha)]\}x_{4t}, \\
\text{s.t.}&\begin{cases} x_{1t} + x_{2t} + x_{3t} + x_{4t} = 1, \\ x_{it} \geqslant 0, i = 1, 2, 3, 4. \end{cases}
\end{aligned}
$$

表 9.3.2 4 种风险证券的投资份额

α	0	0	0	0.3	0.3	0.3	0.5	
β	0	0.5	1	0	0.5	1	0	
投资比例	(0,0,1,0)	(0,0,1,0)	(0,0,1,0)	(0,0,1,0)	(0,1,0,0)	(0,1,0,0)	(0,0,1,0)	
目标函数	−0.018	−0.0145	−0.011	−0.0009	0.0233	0.0725	0.008	
α	0.5	0.5	0.7	0.7	0.7	1	1	1
β	0.5	1	0.3	0.7	1	0	0.5	1
投资比例	(0,1,0,0)	(0,1,0,0)	(0,1,0,0)	(0,1,0,0)	(0,1,0,0)	(0,0,1,0)	(0,1,0,0)	(0,1,0,0)
目标函数	0.0525	0.1308	0.0387	0.1247	0.1892	0.0339	0.1255	0.2767

由表 9.3.2 可以看出, 收益与风险的权衡最大值分别是 α 和 β 的增函数. α 越大, β 越大, 则目标函数值就越大; 反之, α 越小, β 越小, 则目标函数值就越小.

本节通过引入投资者对投资收益和投资风险关心程度参数 α, 将证券组合投资的多目标区间数线性规划模型转化为单目标区间数线性规划模型, 然后通过引入优化水平参数 β, 将目标函数为区间数的区间数线性规划模型转化为一般的参数线性

规划模型来求解. 投资者可以根据对收益风险的关心程度和对风险的喜好程度以及金融市场的客观情况, 适当估计参数 α 和 β, 从而得到相应情况下的有效投资方案, 使投资过程更具柔性, 而且更接近于实际情况.

9.4　剩余劳动力配置的结构模型研究

世界上发展中国家一般具有如下两个基本特征：一是人口众多, 劳动力利用不充分, 这是人均收入较低、劳动生产率低下的一个原因; 二是经济结构具有二元性, 即具有传统的农业生产, 同时具有现代化的工业大生产. 我国是一个典型的发展中国家, 人口资源最多, 人力资源极为丰富. 据有关部门统计, 目前我国城乡剩余劳动力约 1.5 亿人. 因此, 如何合理适度地转移一部分剩余劳动力, 对改变我国二元经济结构, 缩小城乡差别, 实现农村现代化, 完成城市化过程将起到至关重要的作用. 同时也是关系到国有企业改革成败的关键. 本节利用证券组合投资的理论, 采用剩余劳动力转移的收益率的均值和方差作为剩余劳动力转移的两个评价指标, 建立了多目标规划模型.

9.4.1　剩余劳动力转移结构的合理性分析[127]

剩余劳动力的合理有序流动对国民经济的发展起着重要的积极作用. 例如, 传统的农业部门劳动力非常丰富, 根据边际生产率递减的规律, 农业部门劳动生产率较低, 甚至出现农业边际生产率为零的现象. 因此, 劳动力从生产率较低的农业部门向第二产业或第三产业转移, 不仅能提高农业部门劳动生产率, 还有利于提高其他产业部门劳动生产率. 但是剩余劳动力的转移存在一个合理的规模和结构问题.

所谓的规模问题是指剩余劳动力向各个产业部门的转移的总量必须与其对剩余劳动力的需求量相一致. 目前我国剩余劳动力转移属于绝对剩余劳动力的转移的阶段. 因为从某个时点上观察, 剩余劳动力的供给量 l_s 与各产业部门追加劳动力时对剩余劳动力的需求量 l_d 不一致, 表现为 $l_s > l_d$, 根据非均衡条件下的短边法则, 合理的剩余劳动力转移的规模 l^* 应为

$$l^* = \min(l_s, l_d) = l_d. \tag{9.4.1}$$

除了合理的转移规模以外, 剩余劳动力的转移还存在一个合理的结构问题. 所谓的结构问题是指剩余劳动力转移在地区结构和产业结构等方面的平衡关系. 地区结构的平衡关系并不要求剩余劳动力转移在全国范围内是均匀的. 因为各地区剩余劳动力的密度不同, 我们应根据各地区对剩余劳动力的需求情况来确定剩余劳动力的合理有序的流动. 产业结构的平衡关系要求剩余劳动力转移应符合国家的产业政策. 这就是说剩余劳动力的转移方向必须符合国家的产业政策, 要与各产业

部门内部结构比例相协调. 从而有利于提高国民经济的运行效果, 促进国民经济朝着良性方向循环. 而调整产业结构主要通过调整投资结构来实现的. 一般地说, 第 i 个产业部门投资规模越大, 则对剩余劳动力吸纳能力越强, 用函数可表为

$$l_{di} = f_i(I_i), \tag{9.4.2}$$

其中 $f_i'(I_i) > 0, f_i''(I_i) < 0, f_i$ 为非线性函数, I_i 表示第 i 个产业部门投资总额, l_{di} 表示对剩余劳动力需求量. 但 l_{di} 与 I_i 的函数关系一般不是线性的. 这是由于随着科学技术进步, 资本有机构成的提高, 单位投资所吸收剩余劳动力的数量有所下降. 从理论上存在与产业结构相一致的最优投资结构, 此时各产业部门所需的剩余劳动力的结构被唯一确定.

9.4.2 剩余劳动力转移结构的多目标规划模型[127]

假设社会上存在 n 个产业部门. 每个剩余劳动力对各个产业部门的收益可用其劳动生产率来衡量, 用 m_i 表示, m_i 可视为随机变量. 用 M_i 表示其数学期望. 即

$$M_i = \mathrm{E}m_i. \tag{9.4.3}$$

第 i 个部门收益的方差为

$$\sigma_{ii} = E(m_i - Em_i)^2. \tag{9.4.4}$$

则 σ_{ii} 表示剩余劳动力投入到第 i 部门的风险.

设 l_i 表示投入到第 i 产业的剩余劳动力的数量, 则由于剩余劳动力转移到 n 个产业部门给国民经济带来的总期望收益为

$$M(l_1, l_2, \cdots, l_n) = \sum_{i=1}^{n} l_i M_i. \tag{9.4.5}$$

剩余劳动力的转移给国民经济带来的总风险为

$$\sigma^2(l_1, l_2, \cdots, l_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} l_i l_j \sigma_{ij}, \tag{9.4.6}$$

其中 $\sigma_{ij} = E(m_i - Em_i)(m_j - Em_j)$, 它表示第 i 个产业部门与第 j 个产业部门收益的协方差.

设 I_i 表示第 i 个产业部门投资总额, α_i 表示平均单位投资对剩余劳动力的吸纳量. 要求剩余劳动力的转移满足于供需的短边法则. 则有

$$l_i \leqslant \alpha_i I_i, \quad i = 1, 2, \cdots, n. \tag{9.4.7}$$

显然，投入到第 i 产业的剩余劳动力的数量满足非负性，即

$$l_i \geqslant 0, \quad i = 1, 2, \cdots, n. \tag{9.4.8}$$

从而可得剩余劳动力转移的多目标优化模型

$$\begin{aligned} &\max M(l_1, l_2, \cdots, l_n) = \sum_{i=1}^{n} l_i M_i, \\ &\min \sigma^2(l_1, l_2, \cdots, l_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} l_i l_j \sigma_{ij}, \\ &\text{s.t.} \begin{cases} l_i \leqslant \sigma_i I_i, & i = 1, 2, \cdots, n, \\ l_i \geqslant 0, & i = 1, 2, \cdots, n. \end{cases} \end{aligned} \tag{9.4.9}$$

记

$$\boldsymbol{L} = (l_1, l_2, \cdots, l_n)^{\mathrm{T}}, \quad \boldsymbol{M} = (M_1, M_2, \cdots, M_n)^{\mathrm{T}},$$
$$\boldsymbol{V} = (\sigma_{ij})_{n \times n}, \quad \boldsymbol{A} = (\alpha_1 I_1, \alpha_2 I_2, \cdots, \alpha_n I_n)^{\mathrm{T}}.$$

则模型 (9.4.9) 就转化为矩阵形式

$$\begin{aligned} &\max M(\boldsymbol{L}) = \boldsymbol{M}^{\mathrm{T}} \boldsymbol{L}, \\ &\min \sigma^2(\boldsymbol{L}) = \boldsymbol{L}^{\mathrm{T}} \boldsymbol{V} \boldsymbol{L}, \\ &\text{s.t.} \begin{cases} \boldsymbol{L} \leqslant \mathbf{A}, \\ \boldsymbol{L} \geqslant \mathbf{0}. \end{cases} \end{aligned} \tag{9.4.10}$$

9.4.3 模型的求解

对于上述的剩余劳动力的转移的两目标优化模型问题，可采用线性加权和法. 即对收益 $M(\boldsymbol{L})$ 和风险 $\sigma^2(\boldsymbol{L})$ 分别给以权系数作出新的目标函数

$$U(\boldsymbol{L}) = \lambda \sigma^2(\boldsymbol{L}) - (1 - \lambda) M(\boldsymbol{L}) = \lambda \boldsymbol{L}^{\mathrm{T}} \boldsymbol{V} L - (1 - \lambda) \boldsymbol{M}^{\mathrm{T}} \boldsymbol{L},$$

其中 λ 反映了决策者对剩余劳动力转移到 n 个产业部门给国民经济带来的总收益和风险的相对偏好参数. $\lambda \in [0, 1]$, $\lambda = 1$ 表示极端厌恶风险, $\lambda = 0$ 表示极端无视风险. 则多目标优化模型 (9.4.10) 就转化为单目标优化模型

$$\begin{aligned} &\min U(\boldsymbol{L}) = \lambda \boldsymbol{L}^{\mathrm{T}} \boldsymbol{V} \boldsymbol{L} - (1 - \lambda) \boldsymbol{M}^{\mathrm{T}} \boldsymbol{L}, \\ &\text{s.t.} \begin{cases} \boldsymbol{L} \leqslant \boldsymbol{A}, \\ \boldsymbol{L} \geqslant \mathbf{0}. \end{cases} \end{aligned} \tag{9.4.11}$$

显然, 上述模型为一个凸二次规划问题. 由非线性规划理论知, Kuhn-Tucker 条件不仅是极值点存在的必要条件, 而且也是极值点存在的充分条件. 利用 Kuhn-Tucker 条件可将其表示为如下线性规划问题求解. 则模型 (9.4.11) 的最优解 $(l_1^*, l_2^*, \cdots, l_n^*)^{\mathrm{T}}$ 表示剩余劳动力对各个产业部门的最佳配置量.

目前, 几乎所有的经济学家均意识到, 决定一国经济发展的不仅仅是物质技术, 更为重要的是人力资源. 在分析剩余劳动力转移的合理性之后, 把剩余劳动力转移到各个产业部门作为人力资本投资, 类似于物质资本投资, 也存在收益和风险的问题. 本节建立多目标优化模型, 旨在提高剩余劳动力转移的效率.

第 10 章　组合判断矩阵及相关决策问题

10.1　组合判断矩阵的相容性与一致性关系

层次分析法 (AHP)[128,129] 是一种重要的多属性决策方法, 已在社会经济各个领域都有广泛的应用[130~132]. 层次分析法的关键步骤是由专家给出某一准则下的两两比较的互反判断矩阵, 由此按着某种方法推导出排序权向量. 由于客观事物的复杂性, 一些大型的决策问题需要多名专家参与, 以提高决策的科学性. 因此有必要研究组合判断矩阵的相容性和一致性问题.

文献 [133] 研究了加权几何平均组合判断矩阵的一致性及特征值问题, 给出了群组判断中每个判断矩阵与其相应的加权几何平均组合判断矩阵之间的一致性关系、最大特征值关系不等式等. 文献 [134] 给出了衡量互反判断矩阵相容性的一个指标和相容性的判别准则, 这为群组决策的进一步研究提供了基础. 但是上述两个文献的研究尚存在一些不足: 一是文献 [133] 只研究了组合判断矩阵的一致性, 而一致性和相容性是既存在区别又存在联系的两个概念; 二是文献 [134] 只研究了单个判断矩阵与其特征矩阵的相容性, 包括简单几何平均组合排序向量所构成的特征矩阵在对数相容性意义下与群组判断中单个判断矩阵相容性. 文献 [134] 没有考虑加权几何平均排序向量所构成的特征矩阵与组合判断矩阵的情形. 本节在文献 [133, 134] 的基础上, 研究加权几何平均组合判断矩阵的相容性以及相容性和一致性的关系, 在满意相容性这个较弱的条件下获得了加权几何平均组合排序向量所构成的特征矩阵与组合判断矩阵相容性的一些有益的结果. 这些结果为在群组决策中进一步使用加权几何平均排序向量法提供理论依据.

10.1.1　基本概念[136]

为方便起见, 记 M_{R+} 为所有的 n 阶正互反判断矩阵所构成的集合. 记 $N=\{1,2,\cdots,N\}$. 矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 是 n 阶正互反判断矩阵, 即$\boldsymbol{A}$满足 $a_{ij}>0$, $a_{ii}=1$, $a_{ij}=1/a_{ji}$, $\forall i,\ j\in N$, $i\neq j$. 若 $\forall i,\ j,k\in N$, 有 $a_{ik}a_{kj}=a_{ij}$ 成立, 则称 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为完全一致性正互反判断矩阵.

定义 10.1.1　设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R+}$, $\boldsymbol{\Omega}=(\omega_1,\omega_2,\cdots,\omega_n)^{\mathrm{T}}$ 是$\boldsymbol{A}$的最大特征值 $\lambda_{\max}$ 所对应的非负归一化的特征向量, $\sum\limits_{i=1}^{n}\omega_i=1$, $\omega_i\geqslant 0$, $i\in N$, 令

$$w_{ij}=\omega_i/\omega_j,\quad \forall i,j\in N.$$

则称 n 阶矩阵 $\boldsymbol{W}=(w_{ij})_{n\times n}$ 为$\boldsymbol{A}$所对应的特征矩阵.

定义 10.1.2 设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R^+}$, $\boldsymbol{B}=(b_{ij})_{n\times n}\in M_{R^+}$, 若有 $a_{ij}=b_{ij}$, $\forall i,\ j\in N$. 则称$\boldsymbol{A}$和$\boldsymbol{B}$为完全相容的.

定义 10.1.2 表明, 只有当两个正互反判断矩阵相等时, 它们才完全相容.

定义 10.1.3[129] 设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R^+}$, $\boldsymbol{B}=(b_{ij})_{n\times n}\in M_{R^+}$, 令

$$c_{ij}=a_{ij}b_{ij},\quad \forall i,j\in N.$$

则称 $C=(c_{ij})_{n\times n}$ 为$\boldsymbol{A}$和$\boldsymbol{B}$的 Hadamard 乘积, 记为 $C=\boldsymbol{A}\circ\boldsymbol{B}$.

定义 10.1.4[134] 设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R^+}$, $\boldsymbol{B}=(b_{ij})_{n\times n}\in M_{R^+}$, 令

$$C(\boldsymbol{A},\boldsymbol{B})=\boldsymbol{e}^{\mathrm{T}}\left(\boldsymbol{A}\circ\boldsymbol{B}^{\mathrm{T}}\right)\boldsymbol{e},\quad \mathrm{SI}(\boldsymbol{A},\boldsymbol{B})=\frac{1}{n^2}C(\boldsymbol{A},\boldsymbol{B}),$$

则称 $C(\boldsymbol{A},\boldsymbol{B})$ 为$\boldsymbol{A}$和$\boldsymbol{B}$的相容度, $\mathrm{SI}(\boldsymbol{A},\boldsymbol{B})$ 为$\boldsymbol{A}$和$\boldsymbol{B}$的相容性指标, 其中 $e=(1,1,\cdots,1)^{\mathrm{T}}$.

定义 10.1.5[133] 设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R^+}$, 定义 $\boldsymbol{A}^{\alpha}=(a_{ij}^{\alpha})_{n\times n}$, 其中 α 为任意实数.

定义 10.1.6 设 $\boldsymbol{A}_k=(a_{ij}^{(k)})_{n\times n}\in M_{R^+}$, $k=1,2,\cdots,m$. 设 $\boldsymbol{\Omega}_{(k)}=(\omega_1^{(k)},\omega_2^{(k)},\cdots,\omega_n^{(k)})^{\mathrm{T}}$ 为 $\boldsymbol{A}_k$ 的最大特征值所对应的特征向量, 令

$$\overline{\boldsymbol{A}}=\boldsymbol{A}_1^{\alpha_1}\circ\boldsymbol{A}_2^{\alpha_2}\circ\cdots\circ\boldsymbol{A}_m^{\alpha_m},\tag{10.1.1}$$

$$\overline{\boldsymbol{\Omega}}=\boldsymbol{\Omega}_{(1)}^{\alpha_1}\circ\boldsymbol{\Omega}_{(2)}^{\alpha_2}\circ\cdots\circ\boldsymbol{\Omega}_{(m)}^{\alpha_m}.\tag{10.1.2}$$

则称 $\overline{\boldsymbol{A}}=(\overline{a}_{ij})_{n\times n}$ 为 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 的加权几何平均的组合判断矩阵, 称 $\overline{\boldsymbol{\Omega}}=(\overline{\omega}_1,\overline{\omega}_2,\cdots,\overline{\omega}_n)^{\mathrm{T}}$ 为 $\boldsymbol{\Omega}_{(1)},\boldsymbol{\Omega}_{(2)},\cdots,\boldsymbol{\Omega}_{(m)}$ 的加权几何平均的组合权重向量. 其中 $\boldsymbol{A}_k$ 表示对于某个决策问题第 k 个专家给出的正互反判断矩阵, $\alpha_1,\alpha_2,\cdots,\alpha_m$ 分别是 m 个专家的权重, 满足 $\sum\limits_{k=1}^{n}\alpha_k=1$, $\alpha_k>0$, $k=1,2,\cdots,m$. 显然, α_k 越大表示第 k 个专家更权威.

定义 10.1.7 设 $\overline{\boldsymbol{\Omega}}=(\overline{\omega}_1,\overline{\omega}_2,\cdots,\overline{\omega}_n)^{\mathrm{T}}$ 为 $\boldsymbol{\Omega}_{(1)},\boldsymbol{\Omega}_{(2)},\cdots,\boldsymbol{\Omega}_{(m)}$ 的加权几何平均的组合权重向量, 令

$$g_{ij}=\overline{\omega_i}/\overline{\omega_j},\quad \forall i,j\in N.$$

则称 n 阶矩阵 $\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}=(g_{ij})_{n\times n}$ 为 $\overline{\boldsymbol{\Omega}}$ 所构成的特征矩阵.

10.1.2 组合判断矩阵的相容性与一致性的主要结果[136]

引理 10.1.1[129]　$\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R^+}$, $\boldsymbol{A}$ 是完全一致性互反判断矩阵的充分必要条件是$\boldsymbol{A}$存在唯一的满足归一化条件的属于最大特征值 $\lambda_{\max}=n$ 对应的特征向量 $\boldsymbol{\Omega}=(\omega_1,\omega_2,\cdots,\omega_n)^{\mathrm{T}}$, 即

$$\boldsymbol{A\Omega}=n\boldsymbol{\Omega},\quad a_{ij}=\omega_i/\omega_j,\quad \forall i,j\in N.$$

引理 10.1.2[134]　设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R^+}$, $\boldsymbol{A}$与其特征矩阵 $\boldsymbol{W}=(w_{ij})_{n\times n}$ 完全相容的充分必要条件是$\boldsymbol{A}$为完全一致性矩阵.

定理 10.1.1　若每个判断矩阵 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 均是完全一致性判断矩阵, 则加权几何平均的组合判断矩阵 $\bar{\boldsymbol{A}}$ 与其特征矩阵完全相容.

证明　设 $\boldsymbol{\Omega}_{(k)}=(\omega_1^{(k)},\omega_2^{(k)},\cdots,\omega_n^{(k)})^{\mathrm{T}}$ 为 A_k 的最大特征值所对应的权重向量, 若 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 均是完全一致性矩阵, 由引理 10.1.1 知

$$a_{ij}^{(k)}=\omega_i^{(k)}/\omega_j^{(k)},\quad \forall i,j\in N,\quad k=1,2,\cdots,m. \tag{10.1.3}$$

由式 (10.1.1) 得

$$\bar{a}_{ij}=\prod_{k=1}^{m}\left(a_{ij}^{(k)}\right)^{\alpha_k},\quad \forall i,j=1,2,\cdots,n. \tag{10.1.4}$$

由式 (10.1.2) 得

$$\bar{\omega}_j=\prod_{k=1}^{m}\left(\omega_j^{(k)}\right)^{\alpha_k},\quad \forall j=1,2,\cdots,n. \tag{10.1.5}$$

所以由式 (10.1.3),(10.1.5) 得

$$\begin{aligned}\sum_{j=1}^{n}\overline{a}_{ij}\overline{\omega}_j&=\sum_{j=1}^{n}\prod_{k=1}^{m}(a_{ij}^{(k)})^{a_k}\cdot\prod_{k=1}^{m}(\omega_j^{(k)})^{a_k}=\sum_{j=1}^{n}\prod_{k=1}^{m}(a_{ij}^{(k)}\omega_j^{(k)})^{a_k}\\&=\sum_{j=1}^{n}\prod_{k=1}^{m}(\omega_i^{(k)})^{a_k}=\sum_{j=1}^{n}\overline{\omega}_i=n\overline{\omega}_i,\quad \forall i\in N.\end{aligned} \tag{10.1.6}$$

式 (10.1.6) 写成矩阵的形式即为 $\overline{\boldsymbol{A}}\,\overline{\boldsymbol{\Omega}}=n\overline{\boldsymbol{\Omega}}$, 其中, $\overline{\boldsymbol{A}}=(\overline{a}_{ij})_{n\times n}$, $\overline{\boldsymbol{\Omega}}=(\overline{\omega}_1,\overline{\omega}_2,\cdots,\overline{\omega}_n)^{\mathrm{T}}$, 所以 $\overline{\boldsymbol{\Omega}}$ 为 $\overline{\boldsymbol{A}}$ 属于特征值为 n 的特征向量. 对 $\overline{\boldsymbol{\Omega}}$ 作归一化处理, 令 $\overline{\boldsymbol{\Omega}}^*=\overline{\boldsymbol{\Omega}}/\boldsymbol{e}^{\mathrm{T}}\overline{\boldsymbol{\Omega}}$, 其中 $\boldsymbol{e}=(1,1,\cdots,1)^{\mathrm{T}}$. 显然有

$$\overline{\boldsymbol{A}}\,\overline{\boldsymbol{\Omega}}^*=n\overline{\boldsymbol{\Omega}}^*. \tag{10.1.7}$$

因为 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 是互反判断矩阵, 由定义 10.1.6 知, $\overline{\boldsymbol{A}}$ 也是互反判断矩阵, $\overline{\boldsymbol{\Omega}}^*$ 为 $\overline{\boldsymbol{A}}$ 属于特征值为 n 的特征向量. 由引理 10.1.1 知 $\overline{\boldsymbol{A}}$ 是完全一致性矩阵. 再由引理 10.1.2 知 $\overline{\boldsymbol{A}}$ 与其特征矩阵完全相容. 证毕.

定理 10.1.1 表明, 当各个专家给出的正互反判断矩阵 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 均是完全一致性矩阵, 组合判断矩阵 $\overline{\boldsymbol{A}}$ 才与其特征矩阵完全相容. 定理 10.1.1 给出了相容性和一致性的关系.

一般而言, 各个专家给出的正互反判断矩阵 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 很难满足一致性. 在一致性不满足的条件下, 几何平均的组合权重向量 $\overline{\boldsymbol{\Omega}}=(\overline{\omega_1},\overline{\omega_2},\cdots,\overline{\omega_n})^{\mathrm{T}}$ 的有效性如何? 为此, 需要讨论组合判断矩阵 $\overline{\boldsymbol{A}}$ 和 $\overline{\boldsymbol{\Omega}}$ 所构成的特征矩阵 $\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}$ 的相容性.

现介绍两个引理.

引理 10.1.3[135] 设 $x_k>0$, $\alpha_k>0$, $k=1,2,\cdots,m$, 则有 $\prod\limits_{k=1}^{m}x_k^{\alpha_k}\leqslant\sum\limits_{j=1}^{m}\alpha_k x_k$, 当且仅当 $x_1=x_2=\cdots=x_m$ 时等号成立, 其中 $\sum\limits_{k=1}^{m}\alpha_k=1$.

引理 10.1.4[134] 设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_{R^+}$, $\boldsymbol{W}=(w_{ij})_{n\times n}$ 为$\boldsymbol{A}$对应的特征矩阵. 相容性指标 $\mathrm{SI}(\boldsymbol{A},\ \boldsymbol{W})=\dfrac{\lambda_{\max}}{n}\geqslant 1$, 当 $\mathrm{SI}(A,\ W)\leqslant\mathrm{SI}^*$ 时, 则矩阵$\boldsymbol{A}$与 $\boldsymbol{W}$ 具有满意的相容性. 其中 $\mathrm{SI}^*\leqslant 0.1\dfrac{1}{n}\mathrm{RI}+1$, RI 为 AHP 中随机一致性指标, $\lambda_{\max}$ 为$\boldsymbol{A}$的最大特征值.

当 n 不太大时, 可取 $\mathrm{SI}^*=1.10$ 作为统一的临界值[134].

定理 10.1.2 设 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 是 m 个专家给出的正互反判断矩阵, 组合判断矩阵 $\overline{\boldsymbol{A}}$ 和 $\overline{\Omega}$ 所构成的特征矩阵 $\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}$ 相容度 $C(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}})$ 与各个专家给出的正互反判断矩阵 $\boldsymbol{A}_{(k)}$ 分别与其特征矩阵 $\boldsymbol{W}_{(k)}$ 相容度 $C(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})$ 的关系满足如下不等式

$$C(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}})\leqslant\sum_{k=1}^{m}\alpha_k C(\boldsymbol{A}_k,\boldsymbol{W}_{(k)}),\tag{10.1.8}$$

其中$\alpha_1,\alpha_2,\cdots,\alpha_m$分别是$m$个专家的权重,满足$\sum\limits_{k=1}^{n}\alpha_k=1,\alpha_k>0,k=1,2,\cdots,m$.

证明 由定义 10.1.4、定义 10.1.7 和式 (10.1.4)、式 (10.1.5) 得

$$\begin{aligned}C(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}})=e^{\mathrm{T}}(\overline{\boldsymbol{A}}\circ\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}^{\mathrm{T}})e&=\sum_{i=1}^{n}\sum_{j=1}^{n}\overline{a}_{ij}g_{ji}=\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}(a_{ij}^{(k)})^{\alpha_k}(\overline{\omega}_j/\overline{\omega}_i)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}(a_{ij}^{(k)})^{\alpha_k}\left(\prod_{k=1}^{m}(\omega_j^{(k)})^{\alpha_k}\Big/\prod_{k=1}^{m}(\omega_i^{(k)})^{\alpha_k}\right)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}(a_{ij}^{(k)}\omega_j^{(k)}/\omega_i^{(k)})^{\alpha_k}.\end{aligned}\tag{10.1.9}$$

由引理 10.1.3 知

$$\prod_{k=1}^{m}\left(a_{ij}^{(k)}\omega_j^{(k)}/\omega_i^{(k)}\right)^{\alpha_k}\leqslant\sum_{k=1}^{m}\alpha_k(a_{ij}^{(k)}\omega_j^{(k)}/\omega_i^{(k)}).$$

代入式 (10.1.9) 得

$$
\begin{aligned}
C(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}) &\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{m}\alpha_k(a_{ij}^{(k)}\omega_j^{(k)}/\omega_i^{(k)}) = \sum_{k=1}^{m}\alpha_k\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij}^{(k)}\omega_j^{(k)}/\omega_i^{(k)})\right] \\
&= \sum_{k=1}^{m}\alpha_k \boldsymbol{e}^{\mathrm{T}}(\boldsymbol{A}_k\circ \boldsymbol{W}_{(k)}^{\mathrm{T}})\boldsymbol{e} \\
&= \sum_{k=1}^{m}\alpha_k C(\boldsymbol{A}_k,\boldsymbol{W}_{(k)}).\quad \text{证毕.}
\end{aligned}
$$

定理 10.1.2 表明, 加权几何平均组合判断矩阵 $\overline{\boldsymbol{A}}$ 和加权几何平均组合权向量 $\overline{\boldsymbol{\Omega}}$ 所构成的特征矩阵 $\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}$ 相容度, 不大于各个专家给出的正互反判断矩阵与其特征矩阵相容度的加权算术平均.

由定义 10.1.4 和定理 10.1.2 即得

推论 10.1.1 组合判断矩阵$\overline{\boldsymbol{A}}$和$\overline{\boldsymbol{\Omega}}$所构成的矩阵 $\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}$相容性指标$\mathrm{SI}(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}})$与各个专家给出的正互反判断矩阵$\boldsymbol{A}_{(k)}$分别与其特征矩阵 $\boldsymbol{W}_{(k)}$相容性指标$\mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})$满足如下不等式

$$
\mathrm{SI}(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}) \leqslant \sum_{k=1}^{m}\alpha_k \mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)}). \tag{10.1.10}
$$

由引理 10.1.4 和推论 10.1.1 易得

推论 10.1.2 若 $\mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)}) \leqslant \mathrm{SI}^*, k=1,2,\cdots,m$, 则 $\mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}}) \leqslant \mathrm{SI}^*$.

推论 10.1.2 指出, 当各个专家给出的正互反判断矩阵 $\boldsymbol{A}_{(k)}$ 分别与其特征矩阵 $\boldsymbol{W}_{(k)}$ 具有满意的相容性时, 组合判断矩阵 $\overline{\boldsymbol{A}}$ 和 $\overline{\boldsymbol{\Omega}}$ 所构成的特征矩阵 $\boldsymbol{G}_{\overline{\boldsymbol{\Omega}}}$ 也具有满意的相容性, 从而表明加权几何平均的组合权重向量具有一定的有效性, 这为群组决策中利用该法提供了理论依据.

引理 10.1.5[133] 设 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 是 m 个专家给出的正互反判断矩阵, $\lambda_{\max}^{(k)}$ 为$\boldsymbol{A}_{(k)}$的最大特征值,$k=1,2,\cdots,m$, $\overline{\boldsymbol{A}}$ 为组合判断矩阵,$\alpha_k>0$, $k=1,2,\cdots,m$, $\sum\limits_{k=1}^{m}\alpha_k=1$, $\lambda_{\max}$ 为 $\overline{\boldsymbol{A}}$ 的最大特征值, 则有

$$
\lambda_{\max} \leqslant \sum_{k=1}^{m}\alpha_k\lambda_{\max}^{(k)}. \tag{10.1.11}
$$

定理 10.1.3 组合判断矩阵$\overline{\boldsymbol{A}}$和其特征矩阵$\overline{\boldsymbol{W}}$的相容性指标 $\mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})$ 与各个专家给出的正互反判断矩阵 $\boldsymbol{A}_{(k)}$ 分别与其特征矩阵 $\boldsymbol{W}_{(k)}$ 相容性指标 $\mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})$ 的关系满足

$$
\mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}}) \leqslant \sum_{k=1}^{m}\alpha_k \mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)}). \tag{10.1.12}
$$

证明 由引理 10.1.4 知

$$\mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})=\frac{\lambda_{\max}}{n},\quad \mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})=\frac{\lambda_{\max}^{(k)}}{n},\quad k=1,2,\cdots,m.$$

将其代入到式 (10.1.11) 即得式 (10.1.12) . 证毕.

推论 10.1.3 若 $\mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})\leqslant \mathrm{SI}^*, k=1,2,\cdots,m$, 则 $\mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})\leqslant \mathrm{SI}^*$.

推论 10.1.3 表明, 当各个专家给出的正互反判断矩阵 $\boldsymbol{A}_{(k)}$ 分别与其特征矩阵 $\boldsymbol{W}_{(k)}$ 具有满意的相容性时, 则组合判断矩阵 $\overline{\boldsymbol{A}}$ 与其特征矩阵 $\overline{\boldsymbol{W}}$ 也具有满意的相容性.

10.1.3 应用举例分析[136]

下面以文献 [133] 中的数据来验证本节给出的结果. 设有某个决策问题, 有三位专家参与决策, 分别给出了三个判断矩阵

$$\boldsymbol{A}_1=\begin{bmatrix}1 & 2 & 4\\ 1/2 & 1 & 3\\ 1/4 & 1/3 & 1\end{bmatrix},\quad \boldsymbol{A}_2=\begin{bmatrix}1 & 3 & 5\\ 1/3 & 1 & 2\\ 1/5 & 1/2 & 1\end{bmatrix},\quad \boldsymbol{A}_3=\begin{bmatrix}1 & 4 & 5\\ 1/4 & 1 & 3\\ 1/5 & 1/3 & 1\end{bmatrix}.$$

利用特征根法求得相应的权重为

$$\begin{aligned}\boldsymbol{\Omega}_{(1)}&=(0.5584,0.3196,0.1220)^{\mathrm{T}},\\ \boldsymbol{\Omega}_{(2)}&=(0.6483,0.2297,0.1220)^{\mathrm{T}},\\ \boldsymbol{\Omega}_{(3)}&=(0.6738,0.2255,0.1007)^{\mathrm{T}}.\end{aligned}$$

不妨取 3 个专家的权重相等, 即 $\alpha_1=\alpha_2=\alpha_3=1/3$, 则加权几何平均组合判断矩阵为

$$\overline{\boldsymbol{A}}=\boldsymbol{A}_1^{1/3}\circ\boldsymbol{A}_2^{1/3}\circ\boldsymbol{A}_3^{1/3}=\begin{bmatrix}1.0000 & 2.8845 & 4.6416\\ 0.3467 & 1.0000 & 2.6207\\ 0.2154 & 0.3816 & 1.0000\end{bmatrix},$$

用最大特征根法求得 $\overline{\boldsymbol{A}}$ 的权重为

$$\boldsymbol{\Omega}^*=(0.6285\quad 0.2564\quad 0.1151)^{\mathrm{T}}.$$

加权几何平均的组合权重向量

$$\overline{\boldsymbol{\Omega}}=\boldsymbol{\Omega}_{(1)}^{1/3}\circ\boldsymbol{\Omega}_{(2)}^{1/3}\circ\boldsymbol{\Omega}_{(3)}^{1/3}=(0.6248\quad 0.2549\quad 0.1144)^{\mathrm{T}}.$$

所以由 $\overline{\boldsymbol{\Omega}}$ 所构成的特征矩阵为

$$G_{\overline{\Omega}}=\begin{bmatrix}1.0000 & 2.4512 & 5.4615\\ 0.4080 & 1.0000 & 2.2281\\ 0.1831 & 0.4488 & 1.0000\end{bmatrix}.$$

$\overline{\boldsymbol{A}},\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3$ 分别对应的特征矩阵为

$$\overline{\boldsymbol{W}}=\begin{bmatrix}1.0000 & 2.4512 & 5.4605\\ 0.4080 & 1.0000 & 2.2276\\ 0.1831 & 0.4489 & 1.0000\end{bmatrix},\quad \boldsymbol{W}_{(1)}=\begin{bmatrix}1.0000 & 1.7472 & 4.5770\\ 0.5723 & 1.0000 & 2.6197\\ 0.2185 & 0.3817 & 1.0000\end{bmatrix}.$$

$$\boldsymbol{W}_{(2)}=\begin{bmatrix}1.0000 & 2.8224 & 5.3139\\ 0.3543 & 1.0000 & 1.8828\\ 0.1882 & 0.5311 & 1.0000\end{bmatrix},\quad \boldsymbol{W}_{(3)}=\begin{bmatrix}1.0000 & 2.8224 & 5.3139\\ 0.3543 & 1.0000 & 1.8828\\ 0.1882 & 0.5311 & 1.0000\end{bmatrix}.$$

所以相容度和相容性指标分别为

$$C(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\Omega}})=9.0795,\quad C(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})=9.0793,\quad C(\boldsymbol{A}_1,\boldsymbol{W}_{(1)})=9.0549,$$
$$C(\boldsymbol{A}_2,\boldsymbol{W}_{(2)})=9.0111,\quad C(\boldsymbol{A}_3,\boldsymbol{W}_{(3)})=9.2573,$$
$$\mathrm{SI}(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\Omega}})=1.0088,\quad \mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})=1.0088,\quad \mathrm{SI}(\boldsymbol{A}_1,\boldsymbol{W}_{(1)})=1.0061,$$
$$\mathrm{SI}(\boldsymbol{A}_2,\boldsymbol{W}_{(2)})=1.0012,\quad \mathrm{SI}(\boldsymbol{A}_3,\boldsymbol{W}_{(3)})=1.0286.$$

显然有

$$C(\overline{A},\boldsymbol{G}_{\overline{\Omega}})<\frac{1}{3}\sum_{k=1}^{3}C(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})=9.1078,\quad \mathrm{SI}(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\Omega}})<\frac{1}{3}\sum_{k=1}^{3}\mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})=1.0120,$$
$$C(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})<\frac{1}{3}\sum_{k=1}^{3}C(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})=9.1078,\quad \mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})<\frac{1}{3}\sum_{k=1}^{3}\mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})=1.0120.$$

可见, 定理 10.1.2 和定理 10.1.3 的结论得到了验证.

从计算结果可以看出, $\mathrm{SI}(\boldsymbol{A}_k,\boldsymbol{W}_{(k)})<1.10, k=1,2,3$, 因此可以认为判断矩阵 $\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3$ 与其特征矩阵具有满意相容性, 且 $\mathrm{SI}(\overline{\boldsymbol{A}},\overline{\boldsymbol{W}})<1.10, \mathrm{SI}(\overline{\boldsymbol{A}},\boldsymbol{G}_{\overline{\Omega}})<1.10$, 所以组合判断矩阵 $\overline{\boldsymbol{A}}$ 与其特征矩阵以及组合判断矩阵 $\overline{\boldsymbol{A}}$ 与组合权重构成的特征矩阵 $\boldsymbol{G}_{\overline{\Omega}}$ 均具有满意相容性.

10.2　模糊判断矩阵的相容性研究

文献 [134] 针对互反判断矩阵给出了衡量其相容性的一个指标, 而对模糊判断矩阵相容性的研究目前研究甚少, 文献 [137] 试图给出互反判断矩阵和模糊判断矩

阵相容性的一个通用指标, 但是, 该指标从类型上实际和文献 [134] 均是属于乘积型指标, 且显得不够简洁. 根据模糊判断矩阵的特点, 本节提出相容性的一个加法型指标, 该指标简洁, 易计算和理解, 同时研究了模糊判断矩阵相容性和一致性的关系, 这些结果有助于进一步提高群组决策的可靠性和有效性.

10.2.1 模糊判断矩阵的相容性概念[141]

设有某个多属性决策问题, 其方案集表示为论域 $S=\{S_1,S_2,\cdots,S_n\}$, 记 $N=\{1,2,\cdots,n\}$, 二元对比矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为方案直集 $S\times S$ 上的一个模糊子集, a_{ij} 表示方案 S_i 优于 S_j 的程度.

定义 10.2.1[138] 若矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 满足如下两个条件

$$a_{ii}=0.5,\quad \forall i\in N, \tag{10.2.1}$$

$$a_{ij}+a_{ji}=1,\quad a_{ij}\geqslant 0,\quad \forall i,\ j\in N, i\neq j. \tag{10.2.2}$$

则称$\boldsymbol{A}$为模糊判断矩阵.

根据定义 10.2.1 的式 (10.2.2) 易知

$$\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{E}, \tag{10.2.3}$$

其中 $\boldsymbol{E}$ 表示元素全为 1 的 n 阶矩阵, 即

$$\boldsymbol{E}=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{n\times n},$$

式 (10.2.3) 称为模糊判断矩阵的互补性.

实际上, 定义 10.2.1 表明, 若 $a_{ij}=0.5$, 则方案 S_i 和 S_j 同样重要, 若 $a_{ij}\in[0,\ 0.5)$, 则方案 S_i 比 S_j 次重要, 此时 a_{ij} 越小表示方案 S_i 和 S_j 对比就越不重要. 反之, 若 $a_{ij}\in(0.5,\ 1]$, 则方案 S_i 比 S_j 重要, 此时 a_{ij} 越大表示方案 S_i 和 S_j 对比就越重要.

定义 10.2.2[138] 若模糊判断矩阵$\boldsymbol{A}=(a_{ij})_{n\times n}$ 满足

$$a_{ik}a_{kj}a_{ji}=a_{ij}a_{jk}a_{ki},\quad \forall i,\ j,k\in N, i\neq j\neq k. \tag{10.2.4}$$

则称$\boldsymbol{A}$为完全一致性模糊判断矩阵.

定义 10.2.3 矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 和 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 均为模糊判断矩阵, 若有

$$a_{ij}=b_{ij},\quad \forall i,\ j\in N.$$

则称$\boldsymbol{A}$和$\boldsymbol{B}$为完全相容的. 证毕.

定义 10.2.3 表明, 只有当两个模糊判断矩阵相等时, 它们才完全相容. 此时由式 (10.2.3) 得 $\boldsymbol{A}+\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{E}$. 否则, 若$\boldsymbol{A}$和$\boldsymbol{B}$不完全相容, 则 $\boldsymbol{A}+\boldsymbol{B}^{\mathrm{T}}\neq\boldsymbol{E}$, 可以把 $\boldsymbol{A}+\boldsymbol{B}^{\mathrm{T}}-\boldsymbol{E}$ 作为 $\boldsymbol{A}+\boldsymbol{B}^{\mathrm{T}}$ 与 $\boldsymbol{E}$ 的偏差矩阵, 用它的所有元素的绝对值之和来反映$\boldsymbol{A}$和$\boldsymbol{B}$不相容程度, 取绝对值的目的是防止正负抵消, 为此引进如下定义.

定义 10.2.4 矩阵$\boldsymbol{A}=(a_{ij})_{n\times n}$ 和$\boldsymbol{B}=(b_{ij})_{n\times n}$ 均为模糊判断矩阵, 则称

$$FC(\boldsymbol{A},\boldsymbol{B})=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}+b_{ji}-1| \tag{10.2.5}$$

为$\boldsymbol{A}$和$\boldsymbol{B}$的相容度.

由定义 10.2.1 和定义 10.2.4, 不难证明定理 10.2.1 和定理 10.2.2.

定理 10.2.1 模糊判断矩阵$\boldsymbol{A}$和$\boldsymbol{B}$的相容度 $FC(\boldsymbol{A},\boldsymbol{B})\geqslant 0$, 且$\boldsymbol{A}$和$\boldsymbol{B}$为完全相容的充要条件是 $FC(\boldsymbol{A},\boldsymbol{B})=0$

定理 10.2.2 模糊判断矩阵$\boldsymbol{A}$和$\boldsymbol{B}$的相容度满足如下性质:

(1) 自反性: $\boldsymbol{A}$和$\boldsymbol{A}$为完全相容, 即 $FC(\boldsymbol{A},\boldsymbol{A})=0$.

(2) 对称性: $FC(\boldsymbol{A},\boldsymbol{B})=FC(\boldsymbol{B},\boldsymbol{A})$.

(3) 传递性: 若$\boldsymbol{A}$和$\boldsymbol{B}$为完全相容, $\boldsymbol{B}$和$\boldsymbol{C}$为完全相容, 则$\boldsymbol{A}$和$\boldsymbol{C}$为完全相容的.

定理 10.2.3 模糊判断矩阵$\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$的相容度满足如下三角不等式.

$$FC(\boldsymbol{A},\boldsymbol{C})\leqslant FC(\boldsymbol{A},\boldsymbol{B})+FC(\boldsymbol{B},\boldsymbol{C}). \tag{10.2.6}$$

证明 由式 (10.2.5) 得

$$FC(\boldsymbol{A},\boldsymbol{C})=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}+c_{ji}-1|=\sum_{i=1}^{n}\sum_{j=1}^{n}|(a_{ij}+b_{ji}-1)+(-b_{ji}+c_{ji})|. \tag{10.2.7}$$

因为$\boldsymbol{B}$为模糊判断矩阵, 由式 (10.2.2) 得

$$b_{ji}=1-b_{ij},\quad \forall i,j\in N.$$

将其代入式 (10.2.7) , 所以有

$$\begin{aligned}FC(\boldsymbol{A},\boldsymbol{C})&\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}(|a_{ij}+b_{ji}-1|+|b_{ij}+c_{ji}-1|)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}+b_{ji}-1|+\sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}+c_{ji}-1|\\&=FC(\boldsymbol{A},\boldsymbol{B})+FC(\boldsymbol{B},\boldsymbol{C}).\end{aligned}$$

10.2.2 模糊判断矩阵的相容性与一致性的关系[141]

若 $\boldsymbol{D}=(d_{ij})_{n\times n}$ 是互反判断矩阵, 则可以通过如下转换公式[139]

$$a_{ij}=d_{ij}/(1+d_{ij}),\quad \forall i,j\in N, \tag{10.2.8}$$

可得模糊判断矩阵 $\boldsymbol{A}$. 称 $\boldsymbol{A}$ 为互反判断矩阵 $\boldsymbol{D}$ 所对应的模糊判断矩阵.

若$\boldsymbol{A}=(a_{ij})_{n\times n}$ 是模糊判断矩阵, 则可以通过如下转换公式[139]

$$d_{ij}=a_{ij}/(1-a_{ij}),\quad \forall i,j\in N, \tag{10.2.9}$$

可得互反判断矩阵 $\boldsymbol{D}$, 称 $\boldsymbol{D}$ 为模糊判断矩阵 $\boldsymbol{A}$ 所对应的互反判断矩阵.

为叙述方便, 引进如下两个引理.

引理 10.2.1[139] 若 $\boldsymbol{D}$ 是完全一致性互反判断矩阵, 则 $\boldsymbol{D}$ 所对应的模糊判断矩阵$\boldsymbol{A}$也一定是完全一致性的; 反之, 若$\boldsymbol{A}$是完全一致性模糊判断矩阵, 则$\boldsymbol{A}$所对应的互反判断矩阵 $\boldsymbol{D}$ 也一定是完全一致性的.

引理 10.2.2[129] $\boldsymbol{D}$ 是完全一致性互反判断矩阵, $\boldsymbol{w}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 是 $\boldsymbol{D}$ 固有权重向量, 则

$$d_{ij}=w_i/w_j,\quad \forall i,j\in N. \tag{10.2.10}$$

定义 10.2.5 矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为模糊判断矩阵, $\boldsymbol{w}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 是$\boldsymbol{A}$的权重向量, 其中 $\sum\limits_{i=1}^{n}w_i=1, w_i\geqslant 0, i\in N$. 令

$$w_{ij}=w_i/(w_i+w_j)\,,\quad \forall i,j\in N.$$

则称 n 阶矩阵 $\boldsymbol{W}=(w_{ij})_{n\times n}$ 为$\boldsymbol{A}$的特征矩阵.

定理 10.2.4 模糊判断矩阵$\boldsymbol{A}$与其特征矩阵 $\boldsymbol{W}=(w_{ij})_{n\times n}$ 完全相容的充要条件是$\boldsymbol{A}$为完全一致性的.

证明 必要性. 因为 $\boldsymbol{W}$ 为$\boldsymbol{A}$的特征矩阵, 所以由定义 10.2.5 知 $w_{ii}=0.5, \forall i\in N$. 且

$$w_{ij}+w_{ji}=\frac{w_i}{w_i+w_j}+\frac{w_j}{w_j+w_i}=1,\quad w_{ij}\geqslant 0,\quad \forall i,j\in N,\quad i\neq j.$$

所以特征矩阵 $\boldsymbol{W}$ 为模糊判断矩阵. 又因为

$$\begin{aligned}w_{ik}w_{kj}w_{ji}&=\frac{w_i}{w_i+w_k}\frac{w_k}{w_k+w_j}\frac{w_j}{w_j+w_i}\\&=\frac{w_i}{w_i+w_j}\frac{w_j}{w_j+w_k}\frac{w_k}{w_k+w_i}=w_{ij}w_{jk}w_{ki},\quad \forall i,j,k\in N, i\neq j\neq k.\end{aligned}$$

所以由定义 10.2.2 知 $\boldsymbol{W}$ 是完全一致性的. 由必要性知 $\boldsymbol{A}$ 与 $\boldsymbol{W}$ 完全相容, 所以 $\boldsymbol{A}=\boldsymbol{W}$, 从而 $\boldsymbol{A}$ 为完全一致性的.

充分性. 设 $\boldsymbol{w} = (w_1, w_2, \cdots, w_n)^{\mathrm{T}}$ 是完全一致性模糊判断矩阵 $\boldsymbol{A}$ 的权重向量, 由引理 10.2.1 知, $\boldsymbol{A}$ 所对应的互反判断矩阵 $\boldsymbol{D}$ 是完全一致性的, 所以 $\boldsymbol{w} = (w_1, w_2, \cdots, w_n)^{\mathrm{T}}$ 也是 $\boldsymbol{D}$ 的固有权重向量, 由引理 10.2.2 知

$$d_{ij} = w_i / w_j, \quad \forall i, j \in N.$$

将式 (10.2.10) 代入式 (10.2.8) , 则有

$$a_{ij} = w_i / (w_i + w_j), \quad \forall i, j \in N. \tag{10.2.11}$$

所以 $\boldsymbol{A} = \boldsymbol{W}$, 从而它们完全相容. 证毕.

定理 10.2.4 表明, 模糊判断矩阵 $\boldsymbol{A}$ 与其特征矩阵完全相容性等价于 $\boldsymbol{A}$ 自身的完全一致性. 因此, 当 $\boldsymbol{A}$ 具有满意一致性的时候, 可以认为$\boldsymbol{A}$与其特征矩阵具有满意的相容性. 当然, 模糊判断矩阵一致性和相容性又是存在区别的两个概念. 一致性考察的是模糊判断矩阵自身的特征, 而相容性考察的是两个模糊判断矩阵之间的关系. 只有当考察模糊判断矩阵$\boldsymbol{A}$与其特征矩阵的相容性时, 相容性和一致性才是等价的.

定义 10.2.6　设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为模糊判断矩阵, 令

$$I(\boldsymbol{A}, \boldsymbol{B}) = \frac{1}{n^2} FC(\boldsymbol{A}, \boldsymbol{B}). \tag{10.2.12}$$

则称 $I(\boldsymbol{A}, \boldsymbol{B})$ 为$\boldsymbol{A}$和$\boldsymbol{B}$相容性指标.

由定理 10.2.1 得, $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为完全相容的充要条件是 $I(\boldsymbol{A}, \boldsymbol{B}) = 0$. 由定理 10.2.4 得, $\boldsymbol{A}$ 为完全一致性的充要条件是 $\boldsymbol{A}$ 和其特征矩阵 $\boldsymbol{W}$ 的相容指标 $I(\boldsymbol{A}, \boldsymbol{W}) = 0$.

定理 10.2.5　模糊判断矩阵$\boldsymbol{A}$和$\boldsymbol{B}$的相容性指标 $I(\boldsymbol{A}, \boldsymbol{B})$ 满足如下的估计式

$$0 \leqslant I(\boldsymbol{A}, \boldsymbol{B}) \leqslant \max \{|a_{ij} - b_{ij}|, \forall i, j \in N\}. \tag{10.2.13}$$

证明　由定理 10.2.1 知 $FC(\boldsymbol{A}, \boldsymbol{B}) \geqslant 0$, 所以 $I(\boldsymbol{A}, \boldsymbol{B}) \geqslant 0$, 因为$\boldsymbol{B}$为模糊判断矩阵, 所以

$$b_{ji} = 1 - b_{ij}, \quad \forall i, j \in N.$$

从而有

$$\begin{aligned} FC(\boldsymbol{A}, \boldsymbol{B}) &= \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij} + b_{ji} - 1| = \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij} - b_{ij}| \\ &\leqslant \sum_{i=1}^{n} \sum_{j=1}^{n} \max \{|a_{ij} - b_{ij}|\} = n^2 \max \{|a_{ij} - b_{ij}|\}. \end{aligned}$$

从而定义 10.2.6 知结论成立. 证毕.

定义 10.2.7 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为模糊判断矩阵, 若 $I(\boldsymbol{A},\boldsymbol{B})\leqslant\alpha$, 则称$\boldsymbol{A}$和$\boldsymbol{B}$为满意相容的. 特别地, $\boldsymbol{A}$与其特征矩阵 $\boldsymbol{W}$ 的相容性指标 $I(\boldsymbol{A},\boldsymbol{W})\leqslant\alpha$, 则称$\boldsymbol{A}$为满意一致性的. 其中 α 为相容性指标的临界值.

由定理 10.2.5 知 $\alpha\in[0,\ \max\{|a_{ij}-b_{ij}|,\forall i,\ j\in N\}]$, α 的选取体现了决策者的态度, α 越小表明决策者对模糊判断矩阵的一致性要求越高. 通过计算机模拟, 一般可取 $\alpha=0.1$.

10.2.3 基于相容性的模糊判断矩阵的一致性改进方法[142]

定义 10.2.8[143] 令

$$\boldsymbol{T}=\boldsymbol{Q}\dot{+}\boldsymbol{Q}^2\dot{+}\cdots\dot{+}\boldsymbol{Q}^n.$$

则称 $\boldsymbol{T}$ 为模糊判断矩阵 $\boldsymbol{A}$ 的可达矩阵, 其中 $\dot{+}$ 表示布尔运算的“和”. 布尔运算的规则定义为 $0\dot{+}0=0,0\dot{+}1=1,1\dot{+}0=1,1\dot{+}1=1.\boldsymbol{Q}=(q_{ij})_{n\times n}$ 的元素定义为

$$q_{ij}=\begin{cases}1, & \text{当}a_{ij}\geqslant 0.5,i\neq j,\\ 0, & \text{当}a_{ij}<0.5,i\neq j,\\ 0, & \text{当}a_{ij}=0.5,i\neq j.\end{cases}$$

定义 10.2.9[143] 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为模糊判断矩阵, 若$\boldsymbol{A}$中存在 m 个不小于 0.5 的元素, 满足

$$s_{u_1}\succ\cdots\succ s_{u_i}\succ s_{u_j}\succ\cdots\succ s_{u_m}\succ s_{u_1},\quad u_i\in N,\quad i=1,2,\cdots,m.$$

并且在这个方案优劣循环链中至少有一个元素大于 0.5, 那么称$\boldsymbol{A}$是不一致的, 否则称矩阵$\boldsymbol{A}$具有满意一致性. 其中序关系 “$\succ$“表示一个方案优于另一个方案.

定理 10.2.6[143] 若模糊判断矩阵 $\boldsymbol{A}$ 的可达矩阵 $\boldsymbol{T}$ 的对角线上存在为 1 的元素, 则模糊判断矩阵 $\boldsymbol{A}$ 是不一致的, 否则 $\boldsymbol{A}$ 具有满意一致性.

定义 10.2.10[143] 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为模糊判断矩阵, $\boldsymbol{W}=(w_{ij})_{n\times n}$ 为$\boldsymbol{A}$的特征矩阵, 令

$$g_{ij}=a_{ij}-w_{ij},\quad \forall i,j\in N.$$

则称 $\boldsymbol{G}=(g_{ij})_{n\times n}$ 为模糊判断矩阵 $\boldsymbol{A}$ 的偏差矩阵.

由于 $\boldsymbol{A}$ 和 $\boldsymbol{W}$ 均是模糊判断矩阵, 显然可得偏差矩阵 $\boldsymbol{G}$ 是反对称矩阵. 即 $\boldsymbol{G}^{\mathrm{T}}=-\boldsymbol{G}$.

推论 10.2.1 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为模糊判断矩阵, $\boldsymbol{A}$为完全一致性的充要条件是$\boldsymbol{A}$的偏差矩阵 $\boldsymbol{G}=\boldsymbol{0}$.

推论 10.2.2 模糊判断矩阵$\boldsymbol{A}$与其特征矩阵$\boldsymbol{W}$的相容度为

$$FC(\boldsymbol{A},\boldsymbol{W})=\sum_{i=1}^{n}\sum_{j=1}^{n}|g_{ij}|,$$

其中 g_{ij} 是 $\boldsymbol{A}$ 的偏差矩阵 $\boldsymbol{G}$ 的元素.

证明
$$FC(\boldsymbol{A},\boldsymbol{W})=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}+w_{ji}-1|=\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}+\frac{w_j}{w_j+w_i}-1\right|$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}-\frac{w_i}{w_j+w_i}\right|=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}-w_{ij}|=\sum_{i=1}^{n}\sum_{j=1}^{n}|g_{ij}|.$$

由推论 10.2.1 可知, 若偏差矩阵 $\boldsymbol{G}$ 中存在非零元素, 则模糊判断矩阵$\boldsymbol{A}$不具有完全一致性, 并且 $|g_{ij}|$ 偏离 0 越大, 说明对应的 a_{ij} 对$\boldsymbol{A}$的不一致性影响越大, 设 $|g_{i_0j_0}|=\max\limits_{1\leqslant i,j\leqslant n}|g_{ij}|$, 则 $a_{i_0j_0}$ 就是要调整的关键元素. 当 $a_{i_0j_0}>0.5$ 时, 将其适当调小一点, 反之, 当 $a_{i_0j_0}<0.5$ 时, 将其适当调大一点. 一般取调整量 $\lambda=0.05$[143], 直到达到满意一致性时为止. 基于相容性的模糊判断矩阵一致性改进步骤如下:

第一步 计算模糊判断矩阵 $\boldsymbol{A}$ 的可达矩阵 $\boldsymbol{T}$, 由定理 10.2.6 判别 $\boldsymbol{A}$ 是否具有满意一致性, 若是转第五步, 否则转下一步.

第二步 利用特征值法计算 $\boldsymbol{A}$ 所对应的互反判断矩阵的 $\boldsymbol{D}$ 的归一化权向量 $\boldsymbol{w}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$, 从而计算 $\boldsymbol{A}$ 的特征矩阵 $\boldsymbol{W}$.

第三步 计算 $\boldsymbol{A}$ 的偏差矩阵 $\boldsymbol{G}$, 从而确定$\boldsymbol{A}$中要调整的关键元素 $a_{i_0j_0}$.

第四步 若 $a_{i_0j_0}>0.5$, 令 $a'_{i_0j_0}=a'_{i_0j_0}-\lambda, a'_{j_0i_0}=a_{i_0j_0}+\lambda$, 若 $a'_{i_0j_0}<0.5$, 令 $a'_{i_0j_0}=a_{i_0j_0}+\lambda, a'_{i_0j_0}=a_{j_0i_0}-\lambda$, 其他元素不变, 即令 $a'_{i_0j_0}=a_{i_0j_0}, i,j\in N, i\neq i_0, j\neq j_0$, 将矩阵 $\boldsymbol{A}'=(a'_{i_0j_0})_{n\times n}$.

第五步 结束.

10.2.4 应用举例分析[142]

例 10.2.1 下面以文献 [143] 中的数据, 给出了本节提出的基于相容性的模糊判断矩阵一致性改进方法的分析结果. 设有某个决策问题, 其方案集表示为 $X=\{x_1,x_2,x_3,x_4\}$, 决策者给出的模糊判断矩阵为

$$\boldsymbol{A}_0=\begin{bmatrix}0.5 & 0.1 & 0.6 & 0.7\\ 0.9 & 0.5 & 0.8 & 0.4\\ 0.4 & 0.2 & 0.5 & 0.9\\ 0.3 & 0.6 & 0.1 & 0.5\end{bmatrix}.$$

对计算过程作出解释. 先计算模糊判断矩阵 $\boldsymbol{A}_0$ 的可达矩阵 $\boldsymbol{T}_0$, $\boldsymbol{A}_0$ 不具有满意一致性, 因此, 利用特征值法计算 $\boldsymbol{A}_0$ 所对应的互反判断矩阵的 $\boldsymbol{D}$ 的权向量 $\boldsymbol{w}_0$, 从而计算 $\boldsymbol{A}_0$ 的特征矩阵 $\boldsymbol{W}_0$, 然后计算 $\boldsymbol{A}_0$ 的偏差矩阵 $\boldsymbol{G}_0$, 从而确定 $\boldsymbol{A}_0$ 中要调整的关键元素为 a_{42}, 取调整量 λ=0.05, 并将原来的 $a_{42}=0.6$ 调整为 $a'_{42}=0.6-0.05=0.55$, 相应地 $a'_{24}=0.4+0.05=0.45$, 其他元素不变. 调整后的判断矩阵记为 $\boldsymbol{A}_1$, 进入下一轮运算. 即计算 $\boldsymbol{A}_1$ 的可达矩阵 $\boldsymbol{T}_1$, 由此可知 $\boldsymbol{A}_1$ 还不是满意一致性的, 因此, 要

计算出 $\boldsymbol{A}_1$ 的特征矩阵 $\boldsymbol{W}_1$ 和偏差矩阵 $\boldsymbol{G}_1$, 由 $\boldsymbol{G}_1$ 知 $\boldsymbol{A}_1$ 中要调整的关键元素为 a'_{42}, 可将 a'_{42} 再调整为 $a''_{42}=0.5$, 同时将 a'_{24} 调整为 $a''_{24}=0.5$, 其他元素不变, 则调整后的判断矩阵记为 $\boldsymbol{A}_2$. 如此循环, 经过共四步的计算, 最后使调整后的判断矩阵 $\boldsymbol{A}_3$ 所对应的可达矩阵 $\boldsymbol{T}_3$ 的对角线上元素均为 0, 从而 $\boldsymbol{A}_3$ 具有满意的一致性. 具体计算结果见表 10.2.1.

表 10.2.1 模糊判断矩阵一致性改进的计算过程

判断矩阵$\boldsymbol{A}$	可达矩阵	权向量
$\boldsymbol{A}_0=\begin{bmatrix}0.5&0.1&0.6&0.7\\0.9&0.5&0.8&0.4\\0.4&0.2&0.5&0.9\\0.3&0.6&0.1&0.5\end{bmatrix}$	$\boldsymbol{T}_0=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\\1&1&1\end{bmatrix}$	$\boldsymbol{w}_0=\begin{bmatrix}0.1427\\0.4488\\0.2685\\0.1399\end{bmatrix}$
$\boldsymbol{A}_1=\begin{bmatrix}0.5&0.1&0.6&0.7\\0.9&0.5&0.8&0.45\\0.4&0.2&0.5&0.9\\0.3&0.55&0.1&0.5\end{bmatrix}$	$\boldsymbol{T}_1=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\\1&1&1\end{bmatrix}$	$\boldsymbol{w}_1=\begin{bmatrix}0.1427\\0.4684\\0.2615\\0.1275\end{bmatrix}$
$\boldsymbol{A}_2=\begin{bmatrix}0.5&0.1&0.6&0.7\\0.9&0.5&0.8&0.5\\0.4&0.2&0.5&0.9\\0.3&0.5&0.1&0.5\end{bmatrix}$	$\boldsymbol{T}_2=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\\1&1&1\end{bmatrix}$	$\boldsymbol{w}_2=\begin{bmatrix}0.1425\\0.4870\\0.2544\\0.1162\end{bmatrix}$
$\boldsymbol{A}_3=\begin{bmatrix}0.5&0.1&0.6&0.7\\0.9&0.5&0.8&0.55\\0.4&0.2&0.5&0.9\\0.3&0.45&0.1&0.5\end{bmatrix}$	$\boldsymbol{T}_3=\begin{bmatrix}0&0&1&1\\1&0&1&1\\0&0&0&1\\0&0&0&1\end{bmatrix}$	$\boldsymbol{w}_3=\begin{bmatrix}0.1420\\0.5056\\0.2472\\0.1058\end{bmatrix}$

特征矩阵$\boldsymbol{A}$	偏差矩阵
$\boldsymbol{W}_0=\begin{bmatrix}0.5000&0.2413&0.3470&0.5050\\0.7587&0.5000&0.6257&0.7624\\0.6530&0.3743&0.5000&0.6574\\0.4950&0.2376&0.3426&0.5000\end{bmatrix}$	$\boldsymbol{G}_0=\begin{bmatrix}0&-0.1413&0.2530&0.1950\\0.1413&0&0.1743&-0.3624\\-0.2530&-0.1743&0&0.2426\\-0.1950&0.3624&-0.2426&0\end{bmatrix}$
$\boldsymbol{W}_1=\begin{bmatrix}0.5000&0.2335&0.3530&0.5281\\0.7665&0.5000&0.6417&0.7860\\0.6470&0.3583&0.5000&0.6722\\0.4719&0.2140&0.3278&0.5000\end{bmatrix}$	$\boldsymbol{G}_1=\begin{bmatrix}0&-0.1335&0.2470&0.1719\\0.1335&0&0.1583&-0.3360\\-0.2470&-0.1583&0&0.2278\\-0.1719&0.3360&-0.2278&0\end{bmatrix}$
$\boldsymbol{W}_2=\begin{bmatrix}0.5000&0.2264&0.3590&0.5508\\0.7736&0.5000&0.6569&0.8074\\0.6410&0.3431&0.5000&0.6865\\0.4492&0.1926&0.3135&0.5000\end{bmatrix}$	$\boldsymbol{G}_2=\begin{bmatrix}0&-0.1264&0.2410&0.1492\\0.1264&0&0.1431&-0.3074\\-0.2410&-0.1431&0&0.2135\\-0.1492&0.3074&-0.2135&0\end{bmatrix}$
$\boldsymbol{W}_3=\begin{bmatrix}0.5000&0.2195&0.3649&0.5730\\0.7805&0.5000&0.6714&0.8268\\0.6351&0.3286&0.5000&0.7003\\0.4270&0.1732&0.2997&0.5000\end{bmatrix}$	$\boldsymbol{G}_3=\begin{bmatrix}0&-0.1195&0.2351&0.1270\\0.1195&0&0.1286&-0.2768\\-0.2351&-0.1286&0&0.1997\\-0.1270&0.2768&-0.1997&0\end{bmatrix}$

可见, 最终计算结果和文献 [143] 是相同的, 这就说明了本节提出的改进方法的有效性, 同时在一次循环迭代运算中, 本节只需计算一次模糊判断矩阵的特征矩阵即可确定要调整的关键元素, 而文献 [143] 需要计算 n 次以模糊判断矩阵每一行元素为依据完全一致性矩阵, 因此本节的计算较简单些.

例 10.2.2[141]　下面以文献 [137] 中的数据, 给出模糊判断矩阵相容性指标的分析结果. 设有某个决策问题, 其方案集表示为 $X=\{X_1,X_2,X_3\}$, 其中有两位专家给出模糊判断矩阵 $\boldsymbol{B}_1$ 和 $\boldsymbol{B}_2$ 分别为

$$\boldsymbol{B}_1=\begin{bmatrix}0.5 & 0.6 & 0.7\\ 0.4 & 0.5 & 0.6\\ 0.3 & 0.4 & 0.5\end{bmatrix},\quad \boldsymbol{B}_2=\begin{bmatrix}0.5 & 0.7 & 0.8\\ 0.3 & 0.5 & 0.7\\ 0.2 & 0.3 & 0.5\end{bmatrix}.$$

利用文献 [140] 中给出的求解模糊模糊判断矩阵权重公式

$$w_i=\frac{\sum\limits_{j=1}^{n}b_{ij}+\dfrac{n}{2}-1}{n(n-1)},\quad i\in N. \tag{10.2.14}$$

分别求得 $\boldsymbol{B}_1$ 和 $\boldsymbol{B}_2$ 的权重为

$$\boldsymbol{w}_1=(0.3833,0.3333,0.2833)^{\mathrm{T}},\quad \boldsymbol{w}_2=(0.4167,0.3333,0.2500)^{\mathrm{T}}.$$

把 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$ 分别代入式 (10.2.11) 中, 可以计算出模糊判断矩阵 $\boldsymbol{B}_1$ 和 $\boldsymbol{B}_2$ 的特征矩阵 $\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$ 为

$$\boldsymbol{W}_1=\begin{bmatrix}0.5 & 0.535 & 0.575\\ 0.465 & 0.5 & 0.540\\ 0.425 & 0.460 & 0.5\end{bmatrix},\quad \boldsymbol{W}_2=\begin{bmatrix}0.5 & 0.556 & 0.625\\ 0.445 & 0.5 & 0.571\\ 0.375 & 0.429 & 0.5\end{bmatrix}.$$

由式 (10.2.12) 算得模糊判断矩阵 $\boldsymbol{B}_1,\boldsymbol{B}_2$ 以及它们的特征矩阵相容性指标为

$$I(\boldsymbol{B}_1,\boldsymbol{B}_2)=0.0667,\quad I(\boldsymbol{B}_1,\boldsymbol{W}_1)=0.05563,\quad I(\boldsymbol{B}_2,\boldsymbol{W}_2)=0.0997.$$

若取 $\alpha=0.1$, 则有 $I(\boldsymbol{B}_1,\boldsymbol{B}_2)<\alpha$, 可以认为模糊判断矩阵 $\boldsymbol{B}_1$ 和 $\boldsymbol{B}_2$ 满足满意相容性, 同时 $I(\boldsymbol{B}_1,\boldsymbol{W}_1)<\alpha$, $I(\boldsymbol{B}_2,\boldsymbol{W}_2)<\alpha$, 所以 $\boldsymbol{B}_1$ 和 $\boldsymbol{B}_2$ 为满意一致性的.

可见, 本节提出的相容性指标所得的判别相容性和一致性结果和文献 [134] 是相同的. 这就说明了本节提出的相容性指标的有效性.

10.3 模糊判断矩阵排序的最小偏差法的性质

有关模糊判断矩阵的排序理论与方法虽然也取得了一些重要进展[144,145], 但排序理论与方法仍需进一步完善. 本节根据互反判断矩阵和模糊判断矩阵的转换公式, 提出模糊判断矩阵排序的最小偏差方法, 并研究了它的一些性质.

10.3.1 基本概念[147]

在多属性决策中, 专家对决策方案进行两两比较. 若按互反型标度进行赋值, 给出互反判断矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 满足 $a_{ij}>0, a_{ii}=1, a_{ij}=1/a_{ji}, i,j=1,2,\cdots,n$. 若按模糊型标度进行赋值, 给出模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$, 满足 $b_{ii}=0.5, b_{ij}+b_{ji}=1, b_{ij}\geqslant 0, i,j=1,2,\cdots,n$.

为方便起见, 记 $N=\{1, 2, \cdots, n\}$.

定义 10.3.1 设 $\boldsymbol{B}$ 是模糊判断矩阵, 如果对任意的 $k\in N$, $b_{ik}\geqslant b_{jk}$, 有 $v_i\geqslant v_j$, 且前者所有等式成立时, 有 $v_i=v_j$, 则称此种排序方法为强条件下保序的, 其中 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是排序方法得出的排序向量.

定义 10.3.2 若 $b_{ij}\geqslant 0.5$, 对任意 $k\in N$, 有 $b_{ik}\geqslant b_{jk}$; 若 $b_{ij}=0.5$, 对任意 $k\in N$, 有 $b_{ik}\geqslant b_{jk}$, 或者有 $b_{ik}\leqslant b_{jk}$, 则称模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为序传递的.

一种排序方法可以看成判断矩阵集合到排序向量集合的映射, 记为 $T(\cdot)$.

定义 10.3.3 设 $T(\cdot)$ 是一种排序方法, $\boldsymbol{B}$是任意一个给定的判断矩阵, $\boldsymbol{V}=T(\boldsymbol{B})$ 为排序向量, 如果对于任一置换矩阵 $\boldsymbol{P}$, 均有 $\boldsymbol{PV}=T(\boldsymbol{PBP}^{\mathrm{T}})$, 则称这种排序方法是置换不变的.

定义 10.3.4 对任意模糊判断矩阵 $\boldsymbol{B}$, 设 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}=T(\boldsymbol{B})$, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}=T(\boldsymbol{B}^{\mathrm{T}})$ 分别是$\boldsymbol{B}$与 $\boldsymbol{B}^{\mathrm{T}}$ 确定的排序向量, 若在不计较一个规范化因子的情况下, 总有 $w_i=1/v_i, i=1,2,\cdots,n$, 则称这种排序方法是对称的.

定义 10.3.5 设 $T(\cdot)$ 是一种排序方法, $\boldsymbol{V}=T(\boldsymbol{B})$ 为排序向量, 如果 $\boldsymbol{B}$ 是完全一致性矩阵, $\boldsymbol{V}$ 必是$\boldsymbol{B}$的固有权重向量, 则称这种排序方法是相容的.

10.3.2 模糊判断矩阵排序的最小偏差法[147]

设 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 是完全一致的模糊判断矩阵, $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是 $\boldsymbol{B}$ 的排序向量, 其中 $\sum\limits_{i=1}^{n}v_i=1, v_i>0, i\in N$, 则有[146]

$$b_{ij}=\frac{v_i}{v_i+v_j},\quad b_{ji}=\frac{v_j}{v_i+v_j},\quad i,j=1,2,\cdots,n.$$

所以有

$$\frac{b_{ij}}{b_{ji}}=\frac{v_i}{v_j},\quad i,j=1,2,\cdots,n. \tag{10.3.1}$$

上式两边对 i 求和, 并注意到 $\sum\limits_{i=1}^{n} v_i = 1$, 则有

$$v_j = 1 \Big/ \sum_{i=1}^{n} (b_{ij}/b_{ji}) \ , \quad j = 1, 2, \cdots, n. \tag{10.3.2}$$

故得排序向量

$$\boldsymbol{V} = \left(1 \Big/ \sum_{i=1}^{n} (b_{i1}/b_{1i}), \cdots, 1 \Big/ \sum_{i=1}^{n} (b_{in}/b_{ni}) \right)^{\mathrm{T}}. \tag{10.3.3}$$

由于决策者在实际决策时所给出的模糊判断矩阵往往是不完全一致的, 所以式 (10.3.1) 一般不成立, 为此引入偏差项

$$f_{ij} = (b_{ij}/b_{ji})(v_j/v_i) + (b_{ji}/b_{ij})(v_i/v_j) - 2, \quad i, j = 1, 2, \cdots, n.$$

显然 $f_{ij} \geqslant 0$, 为此构造偏差函数 $F(\boldsymbol{V}) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} f_{ij}$, 显然, 合理的排序向量 $\boldsymbol{V} = (v_1, v_2, \cdots, v_n)^{\mathrm{T}}$ 应使 $F(\boldsymbol{V})$ 取得最小值, 即有如下最优化问题

$$\begin{aligned} &\min F(\boldsymbol{V}) = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}, \\ &\text{s.t.} \begin{cases} \sum\limits_{i=1}^{n} v_i = 1, \\ v_i > 0, i = 1, 2, \cdots, n. \end{cases} \end{aligned} \tag{10.3.4}$$

称由最优化问题而得到最优排序向量的方法为模糊判断矩阵排序的最小偏差法, 简记为 LDM.

定理 10.3.1 偏差函数 $F(\boldsymbol{V})$ 在 D 空间中有唯一最小值点 $\boldsymbol{V}^*$, 且 $\boldsymbol{V}^*$ 是方程组

$$\sum_{j=1}^{n} (b_{ij}/b_{ji})(v_j/v_i) = \sum_{j=1}^{n} (b_{ji}/b_{ij})(v_i/v_j), \quad i = 1, \cdots, n \tag{10.3.5}$$

在 D 空间中的唯一最优解, 这里 $D = \left\{ \boldsymbol{V} = (v_1, \cdots, v_n)^{\mathrm{T}} | v_i > 0, i = 1, \cdots, n, \sum\limits_{i=1}^{n} v_i = 1 \right\}$.

证明 对任意的 $\boldsymbol{V} = (v_1, v_2, \cdots, v_n)^{\mathrm{T}} \in D$, 因为 $f_{ij} \geqslant 0$, $i, j = 1, 2, \cdots, n$, 所以 $F(\boldsymbol{V}) \geqslant 0$, 设 $a = \inf\{F(\boldsymbol{V}) | \boldsymbol{V} \in D\}$, 由下确界的定义知, 存在向量序列 $\{\boldsymbol{V}(k) \subset D, k = 1, 2, \cdots\}$, 使得

$$\lim_{k \to \infty} F(\mathbf{V}(k)) = a. \tag{10.3.6}$$

由于$\{\boldsymbol{V}(k)\subset D\}$为有界向量序列, 因而存在收敛的子列$\{\boldsymbol{V}(k_i)=(v_1(k_i),\cdots,v_n(k_i))^{\mathrm{T}}\}$, 设其极限为 $\boldsymbol{V}^*=(v_1^*,v_2^*,\cdots,v_n^*)^{\mathrm{T}}$, 即 $\lim\limits_{k_i\to\infty}\boldsymbol{V}(k_i)=\boldsymbol{V}^*$. 显然 $\sum\limits_{i=1}^{n}v_i^*=1, v_i^*\geqslant 0, i\in N$, 下证 $v_i^*>0,\ i=1,2,\cdots,n$.

假设存在 $v_{i_0}^*=0$, 则

$$\begin{aligned}F(\boldsymbol{V}(k_i))&=\sum_{i=1}^{n}\sum_{j=1}^{n}[(b_{ij}/b_{ji})(v_j(k_i)/v_i(k_i))+(b_{ji}/b_{ij})(v_i(k_i)/v_j(k_i))-2]\\&\geqslant(b_{i_0j}/b_{ji_0})(v_j(k_i)/v_{i_0}(k_i))-2n^2\to\infty.\end{aligned}\tag{10.3.7}$$

式 (10.3.7) 与式 (10.3.6) 矛盾! 所以假设不成立, 从而 $\boldsymbol{V}^*\in D$. 因为收敛序列和其任一个子列都有相同的极限, 且注意到函数 F 的连续性, 则有

$$F(\boldsymbol{V}^*)=\lim_{k_i\to\infty}F(\boldsymbol{V}(k_i))=\lim_{k\to\infty}F(\boldsymbol{V}(k))=a,$$

所以 $\boldsymbol{V}^*\in D$ 是最优化问题式 (10.3.4) 的最优解.

构造 Lagrange 函数 $L(\boldsymbol{V},\lambda)=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}f_{ij}+\lambda\left(\sum\limits_{i=1}^{n}v_i-1\right)$, 因为 $\boldsymbol{V}^*\in D$ 为 D 的内点, 所以它满足如下一阶条件得

$$\frac{\partial L(\boldsymbol{V},\lambda)}{\partial v_i}=\sum_{j=1}^{n}\left[(b_{ij}/b_{ji})(-v_j/v_i^2)+(b_{ji}/b_{ij})(1/v_j)\right]+\lambda=0,\quad i=1,2,\cdots,n.\tag{10.3.8}$$

$$\frac{\partial L(\boldsymbol{V},\lambda)}{\partial\lambda}=\sum_{i=1}^{n}v_i-1=0.\tag{10.3.9}$$

式 (10.3.8) 两边同时乘以 v_i, 对 i 求和, 由式 (10.3.9) 并注意到双重求和可以交换次序得

$$\begin{aligned}\lambda&=\sum_{i=1}^{n}\sum_{j=1}^{n}[(b_{ij}/b_{ji})(v_j/v_i)-(b_{ji}/b_{ij})(v_i/v_j)]\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}(b_{ij}/b_{ji})(v_j/v_i)-\sum_{i=1}^{n}\sum_{j=1}^{n}(b_{ij}/b_{ji})(v_j/v_i)=0.\end{aligned}\tag{10.3.10}$$

由式 (10.3.8)、式 (10.3.10) 即知 V^* 满足方程组式 (10.3.5).

下证方程组解的唯一性. 设 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}\in D$, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}\in D$ 是方程组 (10.3.5) 的两个不同的解, 则存在 k, 使 $v_k/w_k=\max\limits_{i}\{v_i/w_i\}$. 即有

$$v_k/w_k\geqslant v_j/w_j,\quad j=1,2,\cdots,n,\quad \text{至少有一个严格的不等号成立}$$

所以

$$v_k \geqslant v_j w_k/w_j, \quad j=1,2,\cdots,n, \quad \text{至少有一个严格的不等号成立.} \tag{10.3.11}$$

由式 (10.3.11) 得

$$\sum_{j=1}^{n}(b_{jk}/b_{kj})(v_k/v_j) > \sum_{j=1}^{n}(b_{jk}/b_{kj})(v_j w_k/w_j v_j) = \sum_{j=1}^{n}(b_{jk}/b_{kj})(w_k/w_j), \tag{10.3.12}$$

$$\sum_{j=1}^{n}(b_{kj}/b_{jk})(v_j/v_k) < \sum_{j=1}^{n}(b_{kj}/b_{jk})(v_j w_j/w_k v_j) = \sum_{j=1}^{n}(b_{kj}/b_{jk})(w_j/w_k). \tag{10.3.13}$$

因为 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是方程组 (10.3.5) 的解, 所以

$$\sum_{j=1}^{n}(b_{jk}/b_{kj})(v_k/v_j) = \sum_{j=1}^{n}(b_{kj}/b_{jk})(v_j/v_k). \tag{10.3.14}$$

所以由式 (10.3.12) ~ (10.3.14) 得

$$\sum_{j=1}^{n}(b_{kj}/b_{jk})(w_j/w_k) > \sum_{j=1}^{n}(b_{jk}/b_{kj})(w_k/w_j). \tag{10.3.15}$$

式 (10.3.15) 与 $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 是方程组 (10.3.4) 的解矛盾. 所以假设不成立, 即结论成立. 证毕.

10.3.3 模糊判断矩阵排序的最小偏差法的性质[147]

定理 10.3.2 模糊判断矩阵排序的最小偏差方法为强条件下保序的.

证明 设 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 是模糊判断矩阵, 设 $b_{ik}\geqslant b_{jk},\ k=1,2,\cdots,n$, 所以 $1-b_{ik}\leqslant 1-b_{jk}$, 注意到 $\boldsymbol{B}$ 是模糊判断矩阵, 即有 $b_{ki}\leqslant b_{kj}$. 因此

$$b_{ik}/b_{ki} \geqslant b_{jk}/b_{kj}, \quad k=1,2,\cdots,n. \tag{10.3.16}$$

式 (10.3.16) 两边同成以 v_k 并对 k 求和得

$$\sum_{k=1}^{n} b_{ik}v_k/b_{ki} \geqslant \sum_{k=1}^{n} b_{jk}v_k/b_{jk}. \tag{10.3.17}$$

设 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是由最小偏差排序法而得的$\boldsymbol{B}$的排序向量, 由定理 10.3.1 知

$$\sum_{k=1}^{n}(b_{ik}/b_{ki})(v_k/v_i) = \sum_{k=1}^{n}(b_{ki}/b_{ik})(v_i/v_k), \tag{10.3.18}$$

$$\sum_{k=1}^{n}(b_{jk}/b_{kj})(v_k/v_j)=\sum_{k=1}^{n}(b_{kj}/b_{jk})(v_j/v_k). \tag{10.3.19}$$

对式 (10.3.18) 和式 (10.3.19) 式两边分别同乘以 v_i 和 v_j 并注意到式 (10.3.17)，则有

$$\sum_{k=1}^{n}(b_{ki}/b_{ik})(v_i^2/v_k)\geqslant\sum_{k=1}^{n}(b_{kj}/b_{jk})(v_j^2/v_k).$$

即

$$v_i^2\sum_{k=1}^{n}(b_{ki}/b_{ik}v_k)\geqslant v_j^2\sum_{k=1}^{n}(b_{kj}/b_{jk}v_k).$$

由式 (10.3.16) 知 $b_{ki}/b_{ik}\leqslant b_{kj}/b_{jk}$, $k=1,2,\cdots,n$. 所以有

$$\sum_{k=1}^{n}(b_{ki}/b_{ik}v_k)\leqslant\sum_{k=1}^{n}(b_{kj}/b_{jk}v_k),$$

因此有 $v_i^2\geqslant v_j^2$,即$v_i\geqslant v_j$. 证毕.

定理 10.3.3 设模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 是序传递的, $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是最小偏差排序法排序向量, 若 $b_{ij}\geqslant 0.5$, 则 $v_i\geqslant v_j$, 若 $b_{ij}\geqslant 0.5$, 则或者 $v_i\geqslant v_j$, 或者 $v_i\leqslant v_j$.

证明 若 $b_{ij}\geqslant 0.5$, 因为模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 是序传递的, 由定义 10.3.2 知, 对任意的 $k\in N$, 有 $b_{ik}\geqslant b_{jk}$, 由定理 10.3.2 的强条件下保序性知 $v_i\geqslant v_j$.

若 $b_{ij}=0.5$, 由序传递的定义 10.3.2 知, 对任意的 $k\in N$, 有 $b_{ik}\geqslant b_{jk}$, 或者有 $b_{ik}\leqslant b_{jk}$, 同样, 由定理 10.3.2 的强条件下保序性知或者 $v_i\geqslant v_j$, 或者 $v_i\leqslant v_j$, 所以结论成立. 证毕.

定理 10.3.4 模糊判断矩阵排序的最小偏差方法具有置换不变性.

证明 设 $\boldsymbol{B}=(b_{ij})_{m\times n}$ 是模糊判断矩阵, $\boldsymbol{P}$ 为置换矩阵, $\boldsymbol{C}=(c_{ij})_{n\times n}=\boldsymbol{PBP}^{\mathrm{T}}$, 设经置换后 $\boldsymbol{B}$ 的第 i 行成了 $\boldsymbol{C}$ 的第 k 行,$\boldsymbol{B}$ 的第 i 列成了 $\boldsymbol{C}$ 的第 k 列, 且令 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}=\mathrm{LDM}(\boldsymbol{B})$,$\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}=\mathrm{LDM}(\boldsymbol{C})$ 分别是 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 由最小偏差排序法确定的排序向量, 由定理 10.3.1 知

$$\sum_{j=1}^{n}(b_{ij}/b_{ji})(v_j/v_i)=\sum_{j=1}^{n}(b_{ji}/b_{ij})(v_i/v_j).$$

因为 $\boldsymbol{C}$ 的第 k 行等于$\boldsymbol{B}$的第 i 行, $\boldsymbol{C}$ 的第 k 列等于$\boldsymbol{B}$的第 i 列, 且由定理 10.3.1 方程组解的唯一性知, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 是由 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 的第 i 个元素和第 k 个元素置换后得到的. 即有 $\boldsymbol{W}=\boldsymbol{PV}=\mathrm{LDM}(\boldsymbol{PBP}^{\mathrm{T}})=\mathrm{LDM}(\boldsymbol{C})$, 因此模糊判断矩阵排序的最小偏差法是置换不变的. 证毕.

定理 10.3.5 模糊判断矩阵排序的最小偏差法具有相容性.

证明　设模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为一致性矩阵, $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是 $\boldsymbol{B}$ 的固有权重向量, 则有

$$b_{ij}=v_i/(v_i+v_j),\quad i,j=1,2,\cdots,n.$$

所以 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是满足方程组 (10.3.5) 的解. 由定理 10.3.1 知 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是由最小偏差法得出的排序向量. 因而模糊判断矩阵排序的最小偏差法具有相容性. 证毕.

定理 10.3.6　模糊判断矩阵排序的最小偏差法具有对称性.

证明　对于模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$, 设 $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是由最小偏差法得出的排序向量, 则由定理 10.3.1 知, $\boldsymbol{V}=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}$ 是方程组 (10.3.5) 的解. 即有

$$\sum_{j=1}^{n}(b_{ij}/b_{ji})(v_j/v_i)=\sum_{j=1}^{n}(b_{ji}/b_{ij})(v_i/v_j).$$

设 $\boldsymbol{C}=\boldsymbol{B}^{\mathrm{T}},w_i=\dfrac{1}{v_i},i=1,2,\cdots,n$, 因此有

$$\sum_{j=1}^{n}(c_{ij}/c_{ji})(w_j/w_i)=\sum_{j=1}^{n}(b_{ji}/b_{ij})(v_i/v_j)=\sum_{j=1}^{n}(b_{ij}/b_{ji})(v_j/v_i)=\sum_{j=1}^{n}(c_{ji}/c_{ij})(w_i/w_j).$$

即在不计较一个规范化因子的情况下, $\boldsymbol{W}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 也是方程组 (10.3.5) 的解, 所以 $\boldsymbol{W}=\mathrm{LDM}(\boldsymbol{B}^{\mathrm{T}})$, 即最小偏差法具有对称性. 证毕.

10.3.4　实例分析[147]

引用文献 [145] 的数据, 有四个备选方案 x_1,x_2,x_3,x_4, 专家给出模糊判断矩阵

$$\boldsymbol{B}=\begin{bmatrix}0.5&0.7&0.6&0.8\\0.3&0.5&0.4&0.6\\0.4&0.6&0.5&0.7\\0.3&0.4&0.3&0.5\end{bmatrix}.$$

利用本节的最小偏差法, 利用 Matlab 最优化工具箱计算排序向量为

$$\boldsymbol{V}=(0.4264,0.1818,0.2786,0.1132)^{\mathrm{T}},$$

故相应方案排序为 $x_1\succ x_3\succ x_2\succ x_4$, 这与文献 [145] 得到的结果一致. 因此, 本节提出的模糊判断矩阵的最小偏差法是有效的.

10.4 两类区间数判断矩阵的一致性

在层次分析法中, 专家对方案进行两两比较时, 由于信息量不足或方案不够完善, 使得专家没有把握对方案的相对重要性程度作出明确判断, 于是产生了不确定型 AHP, 当对方案两两比较采用区间标度时, 相应的判断矩阵以区间形式给出. 文献 [148,149] 引入了区间数判断矩阵的概念, 文献 [150] 引入了区间数模糊判断矩阵的一致性定义, 并给出了区间数互反判断矩阵和区间数模糊判断矩阵的转换公式. 本节基于文献 [150] 给出的区间数互反判断矩阵和区间数模糊判断矩阵的转换公式给出了它们一致性的充要条件, 并研究了一致性区间数模糊判断矩阵的一些优良性质, 为深入研究不确定性理论奠定了基础.

10.4.1 若干概念[150]

定义 10.4.1[148] 记 $a=[a^-,a^+]=\{x|0<a^-\leqslant x\leqslant a^+\}$, 称 a 为一个区间数, 当 $a^+=a^-$ 时, 区间数 a 即为普通的正实数.

两个区间数 $a=[a^-,a^+], b=[b^-,b^+]$ 称为相等的, 当且仅当 $a^-=b^-, a^+=b^+$, 记为 $a=b$. 常用区间数的加法, 减法和数乘运算规则可参见 9.3, 现介绍其他的运算法则.

乘法运算

$$ab=[a^-b^-,a^+b^+],\text{特别地},\lambda a=[\lambda a^-,\lambda a^+],\lambda>0\text{为实数}.$$

除法运算

$$\frac{a}{b}=\left[\frac{a^-}{b^+},\frac{a^+}{b^-}\right],\text{特别地},\frac{1}{b}=\left[\frac{1}{b^+},\frac{1}{b^-}\right].$$

指数运算

$$c^a=[c^{a-},c^{a+}],\quad c>0.$$

对数运算

$$\log_a b=[\log_a b^-,\log_a b^+],\quad a>1,\quad b^+\geqslant b^->0.$$

上述运算满足相应的交换律、结合律、分配律.

以区间数为元素的向量或矩阵称为区间数向量或区间数矩阵. 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为区间数矩阵,$a_{ij}=[a_{ij}^-,a_{ij}^+]$. 记 $\boldsymbol{A}^-=(a_{ij}^-)_{n\times n},\boldsymbol{A}^+=(a_{ij}^+)_{n\times n}$, 并记 $\boldsymbol{A}=[\boldsymbol{A}^-,\boldsymbol{A}^+]$. 同样对区间数向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}},x_i=[x_i^-,x_i^+]$. 记 $\boldsymbol{x}^-=(x_1^-,x_2^-,\cdots,x_n^-)^{\mathrm{T}},\boldsymbol{x}^+=(x_1^+,x_2^+,\cdots,x_n^+)^{\mathrm{T}}$, 并记为 $\boldsymbol{x}=[\boldsymbol{x}^-,\boldsymbol{x}^+]$.

定义 10.4.2[148] 称 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为一个区间数互反判断矩阵, 对 $\forall i,j=1,2,\cdots,n$, 有

(1) $a_{ij} = [a_{ij}^-, a_{ij}^+]$, 且 $1/9 \leqslant a_{ij}^- \leqslant a_{ij}^+ \leqslant 9$.

(2) $a_{ij} = \dfrac{1}{a_{ji}}, i \neq j, a_{ii} = [1, 1]$.

定义 10.4.3[148] 设 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 为一个区间数互反判断矩阵, 若对 $\forall i, j = 1, 2, \cdots, n$, 均有

$$a_{ij}a_{jk} = a_{jj}a_{ik}. \tag{10.4.1}$$

则称 $\boldsymbol{A}$ 为一致性区间互反判断矩阵.

显然, 当 $a_{ij}^- = a_{ij}^+$ 时, 一致性区间互反判断矩阵即为通常的一致性数字互反判断矩阵.

定义 10.4.4[150] 设 $\boldsymbol{B} = (b_{ij})_{n\times n}, b_{ij} = [b_{ij}^-, b_{ij}^+]$ 为一个区间数判断矩阵, 若对 $\forall i, j = 1, 2, \cdots, n$, 有

$$b_{ii} = [0.5, 0.5], \quad b_{ij}^- + b_{ji}^+ = b_{ij}^+ + b_{ji}^- = 1, \quad i \neq l. \tag{10.4.2}$$

则称 $\boldsymbol{B}$ 是区间数模糊判断矩阵.

定义 10.4.5[150] $\boldsymbol{B} = (b_{ij})_{n\times n}$ 为一个区间数模糊判断矩阵, $b_{ij} = [b_{ij}^-, b_{ij}^+]$, 若有

$$b_{ij} + b_{jk} = b_{jj} + b_{ik}, \quad \forall i, j, k = 1, 2, \cdots, n. \tag{10.4.3}$$

则称 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵. 式 (10.4.3) 称为区间数模糊判断矩阵的一致性条件.

显然, 当 $b_{ij}^- = b_{ij}^+$ 时, $\boldsymbol{B}$ 为一致性数字模糊判断矩阵.

10.4.2 两类区间数判断矩阵的一致性主要结果[150]

设 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 为一个区间数互反判断矩阵, 令

$$b_{ij} = \frac{1}{2} + \log_{81} a_{ij}, \quad \forall i, j = 1, 2, \cdots, n. \tag{10.4.4}$$

则称 $\boldsymbol{B} = (b_{ij})_{n\times n}$ 为区间数互反判断矩阵 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 所对应的矩阵. 实际上对矩阵 $\boldsymbol{B}$ 有如下结论.

引理 10.4.1 若 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 为区间数互反判断矩阵, 则通过转换公式 (10.4.4) 所得的区间数判阵 $\boldsymbol{B} = (b_{ij})_{n\times n}$ 一定是区间数模糊判断矩阵.

证明 因 $b_{ij} = \dfrac{1}{2} + \log_{81} a_{ij}, \quad \forall i, j = 1, 2, \cdots, n$, 由区间数运算规则知

$$b_{ij}^- = \frac{1}{2} + \log_{81} a_{ij}^-, \quad b_{ij}^+ = \frac{1}{2} + \log_{81} a_{ij}^+, \quad \forall i, j = 1, 2, \cdots, n.$$

得到

$$b_{ij}^- + b_{ji}^+ = \left(\frac{1}{2} + \log_{81} a_{ij}^-\right) + \left(\frac{1}{2} + \log_{81} a_{ji}^+\right) = 1 + \log_{81} a_{ij}^- a_{ji}^+.$$

由于 $\boldsymbol{A}$ 为区间数互反判断矩阵, 由定义 10.4.2 知 $a_{ij}=\dfrac{1}{a_{ji}}$, 从而有 $a_{ij}^{-}a_{ji}^{+}=1$, 故有

$$b_{ij}^{-}+b_{ji}^{+}=1.$$

同理可证 $b_{ij}^{+}+b_{ji}^{-}=1$. 由式 (10.4.4) 有 $b_{ii}=[0.5,0.5]$, $\forall i=1,2,\cdots,n$. 由定义 10.4.4 知 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为区间数模糊判断矩阵.

定理 10.4.1 若 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为一致性区间数互反判断矩阵, 则 $\boldsymbol{A}$ 所对应的矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为一致性区间数模糊判断矩阵.

证明 由引理 10.4.1 知, $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为区间数模糊判断矩阵. 因为 $b_{ij}=\dfrac{1}{2}+\log_{81}a_{ij}$, 则有

$$a_{ij}=9^{2b_{ij}-1},\quad \forall i,j=1,2,\cdots,n. \tag{10.4.5}$$

又因为 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为一致性区间数互反判断矩阵, 所以由定义 10.4.3, 有

$$a_{ij}a_{jk}=a_{jj}a_{ik},\quad \forall i,j=1,2,\cdots,n. \tag{10.4.6}$$

将式 (10.4.5) 代入式 (10.4.6), 有 $9^{2b_{ij}-1}\cdot 9^{2b_{jk}-1}=9^{2b_{jj}-1}\cdot 9^{2b_{ik}-1}$, 根据区间数运算规则知

$$9^{2b_{ij}^{-}-1}\cdot 9^{2b_{jk}^{-}-1}=9^{2b_{jj}^{-}-1}\cdot 9^{2b_{ik}^{-}-1},\quad 9^{2b_{ij}^{+}-1}\cdot 9^{2b_{jk}^{+}-1}=9^{2b_{jj}^{+}-1}\cdot 9^{2b_{ik}^{+}-1}.$$

于是得到 $b_{ij}^{-}+b_{jk}^{-}=b_{jj}^{-}+b_{ik}^{-}$, $b_{ij}^{+}+b_{jk}^{+}=b_{jj}^{+}+b_{ik}^{+}$, 即 $b_{ij}+b_{jk}=b_{jj}+b_{ik}$. 由定义 10.4.5 知 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵. 证毕.

设 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为区间数模糊判断矩阵, 令

$$a_{ij}=9^{2b_{ij}-1},\quad \forall i,j=1,2,\cdots,n. \tag{10.4.7}$$

则称 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为区间数模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 所对应的矩阵, 对矩阵 $\boldsymbol{A}$ 有如下结论.

引理 10.4.2 若 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为区间数模糊判断矩阵, 则通过转换公式 (10.4.7) 得到的区间判断矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 一定是区间数互反判断矩阵.

证明 因 $a_{ij}=9^{2b_{ij}-1}$, $\forall i,j=1,2,\cdots,n$, 由区间数运算规则知

$$a_{ij}^{-}=9^{2b_{ij}^{-}-1},\quad a_{ij}^{+}=9^{2b_{ij}^{+}-1},\quad \forall i,j=1,2,\cdots,n.$$

故有

$$a_{ij}^{-}a_{ji}^{+}=9^{2b_{ij}^{-}-1}\cdot 9^{2b_{ji}^{+}-1}=9^{2(b_{ij}^{-}+b_{ji}^{+})-2}.$$

因为$\boldsymbol{B}$为区间数模糊判断矩阵, 由定义 10.4.4 知 $b_{ij}^{-}+b_{ji}^{+}=1$, 从而有 $a_{ij}^{-}a_{ji}^{+}=1$. $\forall i,j=1,2,\cdots,n$. 同理可证 $a_{ij}^{+}a_{ji}^{-}=1$. 显然, 由式 (10.4.7) 有 $a_{ii}=[1,1]$. 由定义 10.4.2 知 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为区间数互反判断矩阵. 证毕.

定理 10.4.2 若 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 为一致性区间数模糊判断矩阵, 则 $\boldsymbol{B}$ 对应的区间判断矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为一致性区间数互反判断矩阵.

证明 由引理 10.4.2 知,$\boldsymbol{A}=(a_{ij})_{n\times n}$ 为区间数互反判断矩阵. 因为 $a_{ij}=9^{2b_{ij}-1}$, 得到

$$b_{ij}=\frac{1}{2}+\log_{81}a_{ij},\quad \forall i,j=1,2,\cdots,n. \tag{10.4.8}$$

又因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 所以由定义 10.4.5 有

$$b_{ij}+b_{jk}=b_{jj}+b_{ik},\quad \forall i,j,k=1,2,\cdots,n. \tag{10.4.9}$$

将式 (10.4.8) 代入式 (10.4.9) , 有

$$\left(\frac{1}{2}+\log_{81}a_{ij}\right)+\left(\frac{1}{2}+\log_{81}a_{jk}\right)=\left(\frac{1}{2}+\log_{81}a_{jj}\right)+\left(\frac{1}{2}+\log_{81}a_{ik}\right).$$

由区间数运算规则知

$$\left(\frac{1}{2}+\log_{81}a_{ij}^-\right)+\left(\frac{1}{2}+\log_{81}a_{jk}^-\right)=\left(\frac{1}{2}+\log_{81}a_{jj}^-\right)+\left(\frac{1}{2}+\log_{81}a_{ik}^-\right),$$

$$\left(\frac{1}{2}+\log_{81}a_{ij}^+\right)+\left(\frac{1}{2}+\log_{81}a_{jk}^+\right)=\left(\frac{1}{2}+\log_{81}a_{jj}^+\right)+\left(\frac{1}{2}+\log_{81}a_{ik}^+\right).$$

于是得到 $a_{ij}^-a_{jk}^-=a_{jj}^-a_{ik}^-,a_{ij}^+a_{jk}^+=a_{jj}^+a_{ik}^+$, 即 $a_{ij}a_{jk}=a_{jj}a_{ik}$. 由定义 10.4.3 知 $\boldsymbol{A}$ 为一致性区间数互反判断矩阵. 证毕.

定义 10.4.6 设区间数模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n}$,

(1) 若 $b_{ij}^{\mathrm{T}}=b_{ji},i,j=1,2,\cdots,n$, 则称 $\boldsymbol{B}^{\mathrm{T}}=(b_{ij}^{\mathrm{T}})_{n\times n}$ 是 $\boldsymbol{B}$ 的转置矩阵.

(2) 若 $b_{ij}^C=1-b_{ij},i,j=1,2,\cdots,n$, 则称 $\boldsymbol{B}^C=(b_{ij}^C)_{n\times n}$ 是 $\boldsymbol{B}$ 的余矩阵.

定理 10.4.3 若 $\boldsymbol{B}=(b_{ij})_{n\times n},b_{ij}=[b_{ij}^-,b_{ij}^+]$ 为一致性区间数模糊判断矩阵, 则 $\boldsymbol{B}$ 具有如下性质:

(1) $\boldsymbol{B}$ 的第 i 行和第 i 列元素之和为 n, $\forall i=1,2,\cdots,n$.

(2) $\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{B}^C$ 且为一致性区间数模糊判断矩阵, $\boldsymbol{B}^{\mathrm{T}}$ 和 $\boldsymbol{B}^{\mathrm{C}}$ 分别是 $\boldsymbol{B}$ 的转置矩阵和余矩阵.

(3) 从 $\boldsymbol{B}$ 中划去任意一行及其对应列所得到的子矩阵是一致性模糊判断矩阵.

证明 (1) 令 $d_i=\sum\limits_{j=1}^n b_{ij},e_i=\sum\limits_{j=1}^n b_{ji}$, 则 d_i 表示 $\boldsymbol{B}$ 的第 i 行元素之和, e_i 表示 $\boldsymbol{B}$ 的第 i 列元素之和. 因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知

$$b_{ij}=b_{ik}+b_{kj}-b_{kk},\quad b_{ji}=b_{jk}+b_{ki}-b_{kk},$$

并注意到 $b_{ii}=[0.5,0.5]=0.5$, 所以有

$$\begin{aligned}d_i+e_i&=\sum_{j=1}^{n}b_{ij}+\sum_{j=1}^{n}b_{ji}=\sum_{j=1}^{n}(b_{ik}+b_{kj}-b_{kk})+\sum_{j=1}^{n}(b_{jk}+b_{ki}-b_{kk})\\&=\sum_{j=1}^{n}(b_{ik}+b_{kj}-b_{kk}+b_{jk}+b_{ki}-b_{kk})\\&=\sum_{j=1}^{n}[(b_{ik}+b_{ki})+(b_{kj}+b_{jk})-(b_{kk}+b_{kk})]\\&=\sum_{j=1}^{n}[(b_{ii}+b_{kk})+(b_{jj}+b_{kk})-(b_{kk}+b_{kk})]\\&=\sum_{j=1}^{n}(b_{ii}+b_{jj})=\sum_{j=1}^{n}(0.5+0.5)=n.\end{aligned}$$

(2) 设 $\boldsymbol{B}^{\mathrm{T}}=(b_{ij}^{\mathrm{T}})_{n\times n}$ 为 $\boldsymbol{B}$ 的转置矩阵, $\boldsymbol{B}^{\mathrm{C}}=(b_{ij}^{\mathrm{C}})_{n\times n}$ 为 $\boldsymbol{B}$ 的余矩阵. 由定义 10.4.6 知

$$b_{ij}^{\mathrm{C}}=1-b_{ij},\quad \forall i=1,2,\cdots,n \tag{10.4.10}$$

根据区间数运算规则知

$$b_{ij}^{\mathrm{C}-}=1-b_{ij}^{+}. \tag{10.4.11}$$

由于 $\boldsymbol{B}$ 为区间数模糊判断矩阵, 由定义 10.4.4 知 $b_{ij}^{+}+b_{ji}^{-}=1$, 故 $b_{ij}^{\mathrm{C}-}=b_{ji}^{-}$. 而由 $b_{ij}^{\mathrm{T}}=b_{ji}$ 可以得到 $b_{ij}^{\mathrm{T}-}=b_{ji}^{-}$. 因此 $b_{ij}^{\mathrm{C}-}=b_{ij}^{\mathrm{T}-}$, 同理可以得到 $b_{ij}^{\mathrm{C}+}=b_{ij}^{\mathrm{T}+}$. 所以 $b_{ij}^{\mathrm{C}}=b_{ij}^{\mathrm{T}},\forall i,j=1,2,\cdots,n$. 因而 $\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{C}}$.

由以上证明可知 $b_{ij}^{\mathrm{C}-}=b_{ji}^{-}$, $b_{ij}^{\mathrm{C}+}=b_{ji}^{+}$, 因此有 $b_{ij}^{\mathrm{C}-}+b_{ji}^{\mathrm{C}+}=b_{ji}^{-}+b_{ij}^{+}=1$, 同理有 $b_{ij}^{\mathrm{C}+}+b_{ji}^{\mathrm{C}+}=1$. 由 (10.4.10) 知 $b_{ii}^{\mathrm{C}}=[0.5,0.5],\forall i=1,2,\cdots,n$. 根据定义 10.4.4 知 $\boldsymbol{B}^{\mathrm{C}}$ 为区间数模糊判断矩阵.

因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 所以由定义 10.4.5 有

$$b_{ij}+b_{jk}=b_{jj}+b_{ik},\quad \forall i,j,k=1,2,\cdots,n.$$

根据区间数运算规则知 $b_{ij}^{+}+b_{jk}^{+}=b_{jj}^{+}+b_{ik}^{+}$, 即

$$b_{ij}^{+}=b_{jj}^{+}+b_{ik}^{+}-b_{jk}^{+}. \tag{10.4.12}$$

将式 (10.4.12) 代入式 (10.4.11), 有

$$\begin{aligned}b_{ij}^{\mathrm{C}-}&=1-b_{ij}^{+}=1-[b_{jj}^{+}+b_{ik}^{+}-b_{jk}^{+}]\\&=(1-b_{ik}^{+})-(1-b_{jk}^{+})+(1-b_{jj}^{+})=b_{ik}^{\mathrm{C}-}-b_{jk}^{\mathrm{C}-}+b_{jj}^{\mathrm{C}-},\end{aligned}$$

即

$$b_{ij}^{C-}+b_{jk}^{C-}=b_{ik}^{C-}+b_{jj}^{C-}.$$

同理有

$$b_{ij}^{C+}+b_{jk}^{C+}=b_{ik}^{C+}+b_{jj}^{C+}.$$

因此有 $b_{ij}^{C}+b_{jk}^{C}=b_{ik}^{C}+b_{jj}^{C}$, 且 $b_{ii}^{C}=1-b_{ii}=[0.5,0.5]$, 由定义 10.4.5 知 $\boldsymbol{B}^{C}$ 为一致性区间数模糊判断矩阵. 因为 $\boldsymbol{B}^{T}=\boldsymbol{B}^{C}$, 所以 $\boldsymbol{B}^{T}$ 也是一致性区间数模糊判断矩阵.

(3) 不失一般性, 设删去 $\boldsymbol{B}$ 中的第 t 行和第 t 列, $\forall t=1,2,\cdots,n$, 则新的矩阵为 $\boldsymbol{B}'=(b'_{ij})_{(n-1)\times(n-1)}$, 其中

$$b'_{ij}=\begin{cases} b_{ij}, & i,j<t,\\ b_{i,j+1}, & i<t,j>t,\\ b_{i+1,j}, & i>t,j<t,\\ b_{i+1,j+1}, & i,j>t. \end{cases}$$

我们分为四种情况讨论:

(i) 若 $i,j<t$,$k<t$, 则有 $b'_{ij}=b_{ij}$,$b'_{jk}=b_{jk}$,$b'_{jj}=b_{jj}$,$b'_{ik}=b_{ik}$. 因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知 $b_{ij}+b_{jk}=b_{jj}+b_{ik}$, 所以有 $b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}$.

若 $i,j<t$,$k>t$, 则有 $b'_{ij}=b_{ij}$,$b'_{jk}=b_{j,k+1}$,$b'_{jj}=b_{jj}$,$b'_{ik}=b_{i,k+1}$. 因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知 $b_{ij}+b_{j,k+1}=b_{jj}+b_{i,k+1}$, 所以有 $b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}$.

(ii) 若 $i<t,j>t$,$k<t$, 则有 $b'_{ij}=b_{i,j+1}$, $b'_{jk}=b_{j+1,k}$,$b'_{jj}=b_{j+1,j+1}$,$b'_{ik}=b_{ik}$. 因为 B 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知 $b_{i,j+1}+b_{j+1,k}=b_{j+1,j+1}+b_{ik}$, 所以有 $b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}$.

若 $i<t$,$j>t$,$k>t$, 则有 $b'_{ij}=b_{i,j+1}$, $b'_{jk}=b_{j+1,k+1}$, $b'_{jj}=b_{j+1,j+1}$, $b'_{ik}=b_{i,k+1}$. 因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知

$$b_{i,j+1}+b_{j+1,k+1}=b_{j+1,j+1}+b_{i,k+1}.$$

所以有 $b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}$.

(iii) 若 $i>t,j<t,k<t$, 则有 $b'_{ij}=b_{i+1,j}$, $b'_{jk}=b_{jk}$, $b'_{jj}=b_{jj}$, $b'_{ik}=b_{i+1,k}$. 因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知 $b_{i+1,j}+b_{jk}=b_{jj}+b_{i+1,k}$, 所以有 $b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}$.

若 $i>t,j<t,k>t$, 则有 $b'_{ij}=b_{i+1,j}$, $b'_{jk}=b_{j,k+1}$, $b'_{jj}=b_{jj}$, $b'_{ik}=b_{i+1,k+1}$. 因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知 $b_{i+1,j}+b_{j,k+1}=b_{jj}+b_{i+1,k+1}$, 所以有 $b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}$.

(iv) 若 $i,j>t,k<t$, 则有 $b'_{ij}=b_{i+1,j+1},b'_{jk}=b_{j+1,k},b'_{jj}=b_{j+1,j+1},b'_{ik}=b_{i+1,k}$. 因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知 $b_{i+1,j+1}+b_{j+1,k}=b_{j+1,j+1}+b_{i+1,k}$, 所以有

$$b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}.$$

若 $i,j>t,k>t$, 则有

$$b'_{ij}=b_{i+1,j+1},\quad b'_{jk}=b_{j+1,k+1},\quad b'_{jj}=b_{j+1,j+1},\quad b'_{ik}=b_{i+1,k+1}.$$

因为 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 由定义 10.4.5 知 $b_{i+1,j+1}+b_{j+1,k+1}=b_{j+1,j+1}+b_{i+1,k+1}$, 所以有

$$b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}.$$

因此 $\forall i,j,k=1,2,\cdots,n;k\neq t$, 均有 $b'_{ij}+b'_{jk}=b'_{jj}+b'_{ik}$, 且 $b'_{jj}=[0.5,0.5]$, 所以 $\boldsymbol{B}'$ 是一致性区间数模糊判断矩阵. 证毕.

综上所述, $\boldsymbol{B}'$ 为一致性区间数模糊判断矩阵.

定理 10.4.3 的意义在于当我们得到区间数一致性模糊判断矩阵后, 如果又要删去某一因素, 则不必重新设计区间数一致性模糊判断矩阵, 而只要从原矩阵中删除该因素对应行与对应列即可. 如此获得的降阶 ($n-1$ 阶) 区间数判断矩阵仍是一致的. 这说明区间数一致性模糊判断矩阵具有很好的鲁棒性.

定理 10.4.4 区间数模糊判断矩阵 $\boldsymbol{B}=(b_{ij})_{n\times n},b_{ij}=[b_{ij}^-,b_{ij}^+]$ 是一致性矩阵的充要条件是任意指定某一行和某一列对应元素之和为常数.

证明 必要性. 设 $\boldsymbol{B}$ 为一致性区间数模糊判断矩阵, 对任意指定的第 i 行和第 j 列, 由定义 10.4.4 知 $\forall k,j=1,2,\cdots,n$, 有 $b_{kk}=b_{jj}$. 而由定义 10.4.5 知,$\forall k=1,2,\cdots,n$ 有

$$b_{ik}+b_{kj}=b_{kk}+b_{ij}=b_{jj}+b_{ij}.$$

在上式中, i 和 j 是固定的, 只有 k 是变动的, 故第 i 行和第 j 列对应元素之和为常数.

充分性. 设 $\boldsymbol{B}$ 为区间数模糊判断矩阵. 对任意指定的第 i 行和第 j 列, 设它们对应元素之和为 a, 即对 $\forall k=1,2,\cdots,n$, 有 $b_{ik}+b_{kj}=a$ 成立. 特别地, 当 $k=j$ 时也成立, 即有 $b_{ij}+b_{jj}=a$. 而由定义 10.4.4 知 $\forall k=1,2,\cdots,n$, 有 $b_{kk}=b_{jj}$. 因此得到 $b_{ik}+b_{kj}=b_{ij}+b_{jj}=b_{ij}+b_{kk}$. 由 i 和 j 的任意性及定义 10.4.5 知, $\boldsymbol{B}$ 是一致性区间数模糊判断矩阵. 证毕.

参 考 文 献

[1] 冯文权，茅奇，周毓萍．经济预测与决策技术（第四版）[M]．武汉：武汉大学出版社,2002.

[2] Bates J M, Granger C W J. The Combination of Forecasts [J]. Operational Research Quarterly, 1969, 20(4): 451~468.

[3] Bunn D W. Forecasting with More than one Model [J]. Journal of Forecasting,1989, 8(3): 161~166.

[4] Granger C W J. Combining Forecasts——Twenty Years Later [J]. Journal of Forecasting,1989, 8(3): 167~173 .

[5] Holden K, Peel D A. Unbiasedness, Efficiency and The Combination of Economic Forecasts [J]. Journal of Forecasting, 1989, 8(3): 175~188.

[6] Hallman J, Kamstra M. Combining Algorithms Based on Robust Estimation Techniques and Co-integrating Restrictions [J]. Journal of Forecasting, 1989, 8(3): 189~198.

[7] Anandalingam G, Chen L. Linear Combination of Forecasts: A General Bayesian Model [J]. Journal of Forecasting, 1989, 8(3): 199~214.

[8] Guerard J B, Clemen R T. Collinearity and the Use of Latent Root Regression for Combining GNP Forecasts [J]. Journal of Forecasting, 1989, 8(3): 231~238.

[9] Sessions D N, Chatterjee S. The combining of forecasts using recursive techniques with non~stationary weights [J]. Journal of Forecasting, 1989, 8(3): 239~251.

[10] Gunter S I, Aksu C. N ~Step Combination of Forecasts [J]. Journal of Forecasting, 1989, 8(3): 253~267.

[11] Wall K D, Correia C. A Preference~based Method for Forecast Combination [J]. Journal of Forecasting, 1989, 8(3): 269~292.

[12] Bischoff C W. The Combination of Macroeconomic Forecasts [J]. Journal of Forecasting, 1989, 8(3): 293~314.

[13] Fiores B E, White E M. Subjective versus Objective Combining of Forecasts: an Experiment [J]. Journal of Forecasting, 1989, 8(3): 331~341.

[14] 唐小我．经济预测与决策新方法及其应用研究 [M]．成都：电子科技大学出版社, 1997.

[15] 唐小我, 马永开, 曾勇, 杨桂元．现代组合预测和组合投资决策方法及应用研究 [M]．北京：科学出版社, 2003.

[16] 唐小我, 曹长修, 金德运．组合预测最优加权向量的进一步研究 [J]．预测, 1994, 13(2): 48~49.

[17] 唐小我．最优组合预测方法及其应用 [J]．数理统计与管理, 1992, 11(1): 31~35.

[18] 唐小我．组合预测误差信息矩阵研究 [J]．电子科技大学学报, 1992, 21(4): 448~454.

[19] 唐小我, 曹长修．组合预测方法研究的若干新结果 [J]．预测, 1992, 11(5): 39~46.

[20] 唐小我, 曾勇, 曹长修．变权组合预测模型研究 [J]．预测, 1993, 12(3): 46~48.

[21] 唐小我, 曾勇．组合预测误差校正模型的应用分析 [J]．管理科学学报, 2002, 5 (6): 53~64.

[22] 曾勇, 唐小我, 郑维敏．基于斯坦规则和误差校正的组合预测模型 [J]．管理科学学报, 2001, 4(6): 39~47.

[23] 王应明, 傅国伟．基于不同误差准则和范数的组合预测方法研究 [J]．控制与决策, 1994, 9(1): 20~28.

[24] 王明涛．确定组合预测权系数最优近似解的方法研究 [J]．系统工程理论与实践, 2000, 20 (3): 104~109.

[25] 谢开贵, 周家启．变权组合预测模型研究 [J]．系统工程理论与实践, 2000, 20 (7): 36~40.

[26] 曾勇, 唐小我．几种无偏组合预测模型的分析 [J]．数量经济技术经济研究, 1996, (11): 41~45.

[27] 马永开, 杨桂元, 唐小我．非负权重组合预测的冗余定理 [J]．系统工程理论方法应用, 1995, 4(4): 33~39.

[28] Dickinson J P. Some Comments on the Combination of Forecasts [J]. Operational Research Quarterly, 1975, 26(1): 205~210.

[29] Armstrong J S. Research on Forecasting: A Quarter~Century Review [J]. Interfaces, 1986, 16(1): 89~109.

[30] Bunn D W. Combining Forecasts [J]. European Journal of Operation Research, 1988, 33(3): 223~229.

[31] Clemen R T. Combining Forecasts: A Review and Annotated Bibliography [J]. International Journal of Forecasting, 1989, 5(4): 559~583.

[32] Granger C W J. Ramanathan R., Improved Methods of Combining Forecasts [J]. Journal of Forecasting, 1984, 3(2): 197~204.

[33] Clemen R T. Linear Constrains and the Efficiency of combined Forecasts [J]. Journal of Forecasting, 1986, 5(1): 31~38.

[34] Trenkler G, Liski E P. Note: Linear Constrains and the Efficiency of combined Forecasts [J]. Journal of Forecasting, 1986, 5(3): 197~202.

[35] Bunn D W. A Bayesian Approach to the Linear Combination of Forecasts [J]. Operations Research Quarterly, 1975, 26(2): 325~329.

[36] Clemen R T, Winkler R L. Combining Economic Forecasts [J]. Journal of Business and Economic Statistics, 1986, 4(1): 39~46.

[37] 高仁祥, 张世英, 刘豹. 组合预测贝叶斯方法研究 [J]. 系统工程学报, 1996, 11(1): 28~35.

[38] 王应明, 傅国伟. 群组预测集结方法研究 [J]. 预测, 1993, 12(3): 42~45.

[39] 周传世. 非线性权组合预测模型及其最优权的确定 [J]. 预测, 1994, 13(2): 60~61.

[40] 傅庚, 唐小我, 曾勇. 广义递归方差倒数组合预测方法研究 [J]. 电子科技大学学报, 1995, 24(2): 211~217.

[41] 唐小我, 曹长修. 递归等权组合预测方法研究 [J]. 电子科技大学学报, 1992, 21 (5): 545~550.

[42] X W Tang, Z F Zhou, Y Shi. The error bounds of combined forecasting[J]. Mathematical and computer modeling, 2002,36, 997~1005.

[43] 王应明. 基于相关性的组合预测方法研究 [J]. 预测, 2002, 21(2): 58~62.

[44] 王应明, 罗英. 调和平均组合预测中的参数估计技术 [J]. 系统工程与电子技术, 1997, 19(10): 45~49.

[45] 王应明. 广义加权几何平均组合预测方法研究 [J]. 厦门大学学报 (自然科学版) , 1998, 37(1): 6~10.

[46] 李正龙. 经济预测与决策方法 [M]. 合肥: 安徽大学出版社, 2002.

[47] 李一智等. 经济预测技术 [M]. 北京: 清华大学出版社, 1991.

[48] 冯文权. 经济预测与决策技术 [M]. 武汉: 武汉大学出版社, 1994.

[49] 王松桂. 线性模型的理论及其应用 [M]. 合肥: 安徽教育出版社, 1987.

[50] Box G E P, Jenkins G M, Reinsel G C. 时间序列分析: 预测与控制 [M]. 顾岚主译. 北京: 中国统计出版社, 1997,211~376.

[51] 邓聚龙. 灰理论基础 [M]. 武汉: 华中科技大学出版社, 2002.

[52] 邓聚龙. 灰色系统理论教程 [M]. 武汉: 华中理工大学出版社, 1990.

[53] 陈华友, 吴涛, 许义生. 灰关联空间与灰关联度计算的改进 [J]. 安徽大学学报 (自然科学版), 1999, 23(4): 1~4.

[54] 王建平. 组合预测权重计算公式的探讨 [J]. 预测, 1993, 12 (4): 54~55.

[55] 项静恬, 史久恩. 非线性系统中数据处理的统计方法 [M]. 北京: 科学出版社, 1997.

[56] 陈华友. 组合预测权系数确定的一种合作对策方法 [J]. 预测, 2003, 22 (1): 75~77.

[57] 王建华. 对策论 [M]. 清华大学出版社, 1986.

[58] 陈华友. 熵值法及其在确定组合预测权系数中的应用 [J]. 安徽大学学报 (自然科学版), 2003, 27(4): 1~6.

[59] 孙庆凯. 平均预测法的应用条件 [J]. 预测, 1985, 4(2): 18~21.
[60] 侯定丕, 陈华友, 陈伟. 数量经济分析 [M]. 北京: 科学出版社, 2004.
[61] 胡运权等. 运筹学教程 [M]. 北京: 清华大学出版社, 2003.
[62] 钱颂迪等. 运筹学 [M]. 北京: 清华大学出版社, 1990.
[63] 陈华友. 全距在确定组合预测最优权系数中的应用 [J]. 安徽大学学报 (自然科学版), 2001, 25 (3): 7~10.
[64] 陈希孺. 数理统计引论 [M]. 北京: 科学出版社, 1997.
[65] 唐纪, 王景. 组合预测方法评述 [J]. 预测, 1999, 18 (2): 42~43.
[66] 曾勇, 唐小我, 曹长修. 非负权重最优组合预测方法研究 [J]. 管理工程学报, 1995, 9(3): 153~161.
[67] 马永开, 唐小我, 杨桂元. 优性组合预测方法存在判别定理 [J]. 预测, 1995, 14(2): 57~58.
[68] 马永开, 杨桂元, 唐小我. 非负权重组合预测的冗余定理 [J]. 系统工程理论方法应用, 1995, 4(4): 33~39.
[69] 马永开, 唐小我, 杨桂元. 非负权重最优组合预测方法的基本理论研究 [J]. 运筹与管理, 1997, 6(2): 1~8.
[70] 王明涛. 预测方法有效性指标一般形式初探 [J]. 预测, 1998, 17(2): 39~40.
[71] 王明涛. 预测方法有效性的进一步分析 [J]. 预测, 1997, 16(3): 50~52.
[72] 陈华友. 基于预测有效度的组合预测模型研究 [J]. 预测, 2001, 20(3): 72~73.
[73] 陈华友, 侯定丕. 基于预测有效度的优性组合预测模型研究 [J]. 中国科学技术大学学报, 2002, 32 (2): 172~180.
[74] 陈华友, 侯定丕. 基于标准差的预测有效度的组合预测模型 [J]. 系统工程学报, 2003, 18 (3): 203~210.
[75] 陈华友, 侯定丕. 基于预测有效度的组合预测方法冗余信息的判定 [J]. 数学的实践与认识, 2004, 34(1): 56~64.
[76] 陈华友. 基于预测有效度的组合预测模型求解方法研究 [J]. 安徽大学学报 (自然科学版), 2003, 27(3): 6~11.
[77] 王洪瑞, 王明涛. 王丰. 河南省化工专门人才需求预测及对策研究 [J]. 预测,1994,13 (6): 36~39.
[78] 王明涛. 确定组合预测权系数最优近似解的方法研究 [J]. 统计研究,1999,16 (7): 43~49.
[79] 陈华友, 许义生. 组合预测的权估计及其显著性检验 [J]. 运筹与管理, 2000, 9 (2): 75~78.
[80] 周传世, 罗国民. 加权几何平均组合预测模型及其应用 [J]. 数理统计与管理, 1995, 2,17~19.
[81] 陈华友. 基于 L_1 范数的加权几何平均组合预测方法 [J]. 安徽大学学报 (自然科学版), 2004, 28(4): 5~10.
[82] 杨桂元, 唐小我, 马永开. 最优加权几何平均组合预测方法研究 [J]. 统计研究, 1996, 2,55~58.
[83] 陈华友, 刘春林. 基于 L1 范数的加权几何平均组合预测方法的性质 [J]. 东南大学学报 (自然科学版), 2004, 34(4): 535~540.
[84] 陈华友, 盛昭瀚, 刘春林. 调和平均的组合预测方法之性质研究 [J]. 系统工程学报, 2004, 19(6): 620~624.
[85] 王应明, 罗英. 广义加权算术平均组合预测技术研究 [J]. 预测, 1998, 17 (1): 51~53.
[86] 陈华友. 广义加权算术平均组合预测法的最优化理论基础及性质 [J]. 系统工程理论与实践, 2003, 23(4): 37~41.
[87] Reeves G R, Lawrence K D. Combining Forecasts Given Different Types of Objectives [J]. European Journal of Operational Research, 1991, 51(1): 65~72.
[88] Lam K F, Mui H W, Yuen H K. A note on minimizing absolute percentage error in combined forecasts[J]. Computers & Operations Research, 2001, 28(11): 1141~1147.
[89] 安徽统计局编. 安徽统计年鉴 —2004[M]. 北京: 中国统计出版社, 2004.
[90] 陈华友. 基于相关系数的优性组合预测模型研究 [J]. 系统工程学报, 2006, 21(4): 353~360.
[91] 陈华友, 赵佳宝, 刘春林. 基于灰色关联度的组合预测模型的性质 [J]. 东南大学学报 (自然科学版), 2004, 34(1): 130~134.
[92] 陈华友, 盛昭瀚, 刘春林. 基于向量夹角余弦的组合预测模型的性质研究. 管理科学学报, 2006, 9(2): 1~8.

[93] 陈华友. 基于 Theil 不等系数的优性组合预测模型的性质 [J]. 电子科技大学学报 (自然科学版), 2004, 33(1): 105～108.

[94] Yager R R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking[J]. IEEE Transactions on Systems, Man and Cybernetics, 1988, 18,183 ～190.

[95] Yager R R. Families of OWA operators[J]. Fuzzy Sets and Systems, 1993,59,125 ～148.

[96] Yager R R. Induced aggregation operators[J]. Fuzzy Sets and Systems, 2003, 137,59～69.

[97] Yager R R. Solving mathematical programming problems with OWA operators as objective functions[A] // Proceeding 1995 IEEE International Conference on Fuzzy systems, 1995,1441～1446.

[98] Yager R R. Fuzzy aggregation of modular neural networks with ordered weighted averaging operators[J]. International Journal of Approximate Reasoning, 1995, 13,359～375.

[99] Herrera F, Herrera～Viedma E, Chiclana F. Multiperson decision～making based on multiplicative preference relations[J]. European Journal of Operation Research, 2001, 129,372～385.

[100] 樊治平, 姜艳萍. 基于 OWG 算子的不同形式偏好信息的群决策方法 [J]. 管理科学学报, 2003, 6(1): 32～36.

[101] 王欣荣, 樊治平. 一种具有不同形式偏好信息的群决策方法 [J]. 东北大学学报 (自然科学版) , 2003, 24(2): 178～181.

[102] 徐泽水. 多属性决策中四类偏好信息的一种集成途径 [J]. 系统工程理论与实践, 2002, 22(11): 117～120.

[103] Yager R R, Filev D P, Sadeghi T. Analysis of flexible structured fuzzy logic controllers[J]. IEEE Transactions on Systems, Man and Cybernetics, 1994, 24,1035～1043.

[104] Yager R R, Goldstein L S, Mendels E F. An approach to aggregating market research data based on fuzzy reasoning[J]. Fuzzy Sets and Systems,1994, 68,1～11.

[105] 徐泽水. 不确定多属性决策方法与应用 [M]. 北京: 清华大学出版社, 2005.

[106] Xu Z S, Da Q L. The ordered weighted geometric averaging operators[J]. International Journal of Intelligent Systems, 2002, 17,709～716.

[107] 陈华友, 刘春林, 盛昭瀚. IOWHA 算子及其在组合预测中的应用 [J]. 中国管理科学, 2004, 12(5): 35～40.

[108] 陈华友, 刘春林. 基于 IOWA 算子的组合预测方法 [J]. 预测, 2003, 22(6): 61～65.

[109] 陈华友, 盛昭瀚. 一类基于 IOWGA 算子的组合预测新方法 [J]. 管理工程学报, 2005, 19(4): 36～39.

[110] Filev D, Yager R R. On the issue of obtaining OWA operator weights[J]. Fuzzy Sets and Systems,1998, 94(2): 157～169.

[111] 王应明. 调和平均组合预测方法的进一步研究 [J]. 系统工程与电子技术, 1998, 20(9): 42～45.

[112] 樊治平, 赵萱. 多属性决策中权重确定的主客观赋权法 [J]. 决策与决策支持系统, 1997, 7(4): 87～91.

[113] 樊治平, 张全, 马建. 多属性决策中权重确定的一种集成方法 [J]. 管理科学学报, 1998, 1(3): 50～53.

[114] 王应明. 离差平方和的多指标决策方法及其应用 [J]. 中国软科学, 2000, 3,110～113.

[115] 陈华友. 多属性决策中的一种最优组合赋权方法研究 [J]. 运筹与管理, 2003, 12(2): 6～10.

[116] 陈华友. 多属性决策中基于离差最大化的的组合赋权方法 [J]. 系统工程与电子技术, 2004, 26 (2): 194～197.

[117] 陈华友, 王居平. 证券组合投资的动态模型研究 [J]. 工科数学, 2001, 17(2): 16～20.

[118] Sharpe W F. 证券投资原理 [M]. 杨秀苔, 刘星等编译. 重庆: 重庆大学出版社, 1998.

[119] 路应金, 唐小我, 周宗放. 证券组合投资的区间数线性规划方法 [J]. 系统工程学报, 2004, 19(1): 33～37.

[120] 达庆利, 徐泽水. 不确定多属性决策的单目标最优化模型 [J]. 系统工程学报, 2002, 17(1): 50～55.

[121] 徐泽水. 求解不确定多属性决策问题的一种新方法 [J]. 系统工程学报, 2002, 17(2): 177～181.

[122] 徐泽水, 孙在东. 一类不确定多属性决策问题的排序方法 [J]. 管理科学学报, 2002, 5(3): 35～39.

[123] Tong S. Interval number and fuzzy number linear programming[J]. Fuzzy Sets and Systems, 1994,66,301～306.
[124] 刘新旺, 达庆利. 一种区间数线性规划的满意解 [J]. 系统工程学报, 1999, 14(2): 123～128.
[125] 达庆利, 刘新旺. 区间数线性规划极其满意解 [J]. 系统工程理论与实践, 1999, 19(12): 42～47.
[126] 赵玉梅, 陈华友. 证券组合投资的多目标区间数线性规划模型 [J]. 运筹与管理, 2006, 15(2): 124～127.
[127] 陈华友, 许义生. 剩余劳动力配置的多目标规划模型研究 [J]. 运筹与管理, 2001, 10 (3): 80～82.
[128] Saaty T L. The Analytic Hierarchy Process[M]. New York: McGraw～Hill, 1980.
[129] 王莲芬, 许树柏. 层次分析法引论 [M]. 北京: 中国人民大学出版社, 1990.
[130] Ramanathan R, Ganesh LS. Using AHP for resource allocation problems [J]. European Journal of Operational Research, 1995, 80(2): 410～417.
[131] Poh K L, Ang B W. Transportation fuels and policy for Singapore: an AHP planning approach [J]. Computers & Industrial Engineering,1999, 37(3): 507～525.
[132] Handfield R, Walton S V, Sroufe R. etc. Applying environmental criteria to supplier assessment: A study in the application of the Analytical Hierarchy Process [J]. European Journal of Operational Research, 2002, 141(1): 70～87.
[133] 徐泽水. 综合判断矩阵的一致性及特征值问题研究 [J]. 系统工程学报, 2000, 15(3): 258～261.
[134] 王莲芬. 相容性与群组决策 [J]. 系统工程理论与实践, 2000, 20(2): 92～96.
[135] Xu Z S , Wei C P. A consistency improving method in the analytic hierarchy process[J]. European Journal of Operational Research, 1999, 116(2): 443～449.
[136] 陈华友, 刘春林. 组合判断矩阵的相容性与一致性关系 [J]. 系统工程理论方法应用, 2004, 13(4): 377～380.
[137] 徐泽水. 衡量判断矩阵相容性的一个通用指标 [J]. 东南大学学报 (自然科学版) , 2001, 31(6): 94～97.
[138] 樊治平, 姜艳萍, 肖四汉. 模糊判断矩阵的一致性及其性质 [J]. 控制与决策, 2001, 16(1): 69～71.
[139] 徐泽水, 达庆利. 三种基于互反判断矩阵的互补判断矩阵排序法 [J]. 东南大学学报 (自然科学版) , 2001, 31(5): 106～109.
[140] 徐泽水. 模糊互补判断矩阵排序的一种算法 [J]. 系统工程学报, 2001, 16(4): 311～314.
[141] 陈华友, 赵佳宝. 模糊判断矩阵的相容性研究 [J]. 运筹与管理, 2004, 13 (1): 44～47.
[142] 陈华友, 刘春林. 基于相容性的模糊判断矩阵一致性改进新方法 [J]. 模糊系统与数学, 2005, 19(2): 122～127.
[143] 姜艳萍, 樊治平. 一种校正模糊判断矩阵一致性的新方法 [J]. 模糊系统和数学, 2002, 16(2): 74～78.
[144] 徐泽水. 互补判断矩阵的两种排序方法—权的最小平方法及特征向量法 [J]. 系统工程理论与实践, 2002, 22(7): 71～75.
[145] 孔松泉, 达庆利, 徐泽水. 互补判断矩阵排序的广义卡方法 [J]. 东南大学学报 (自然科学) , 2002, 6(4): 659～662.
[146] 徐泽水. AHP 中两类标度的关系研究 [J]. 系统工程理论与实践, 1999, 19(7): 97～101.
[147] 陈华友, 周礼刚. 互补判断矩阵排序的最小偏差法的性质 [J]. 运筹与管理, 2004, 13(3): 39～43.
[148] 魏毅强, 刘进生, 王绪柱. 不确定型 AHP 中判断矩阵的一致性概念及权重 [J]. 系统工程理论与实践, 1994, 7 (4): 16～22.
[149] 刘心报. 不确定型 AHP 中判断矩阵的一致性定义 [J]. 运筹与管理, 1998, 7 (2): 41～43.
[150] 周礼刚, 陈华友. 两类区间数判断矩阵一致性研究 [J]. 运筹与管理, 2005, 14(4): 47～51.
[151] 陈华友, 陈启明, 李洪岩. 一类基于 OWA 算子的组合预测模型及其性质 [J]. 运筹与管理, 2006, 15(6): 34～39.

《运筹与管理科学丛书》已出版书目

1．非线性优化计算方法　袁亚湘　著　2008 年 2 月
2．博弈论与非线性分析　俞建　著　2008 年 2 月
3．蚁群优化算法　马良等　著　2008 年 2 月
4．组合预测方法有效性理论及其应用　陈华友　著　2008 年 2 月

[General Information]
书名=组合预测方法有效性理论及其应用
作者=陈华友著
页数=257
SS号=11958503
DX号=
出版日期=2008.2
出版社=科学出版社